

**UNIVERSIDAD NACIONAL DE SAN CRISTÓBAL
DE HUAMANGA**

**FACULTAD DE INGENIERÍA DE MINAS, GEOLOGÍA Y
CIVIL**

ESCUELA PROFESIONAL DE INGENIERÍA CIVIL



TESIS:

**Detección de cuerpos de agua mediante el procesamiento de
imágenes satelitales utilizando un modelo de Deep Learning
en la provincia de Cangallo, 2024**

Para optar el título profesional de:

INGENIERO CIVIL

PRESENTADO POR:

Bach. Alex Elven CHAVEZ TITO

ASESOR:

Msc. Ing. Edward LEÓN PALACIOS

AYACUCHO - PERÚ

2026

RESUMEN

La presente investigación evaluó un modelo de Deep Learning basado en la arquitectura U-Net para la segmentación automatizada de cuerpos de agua en la provincia de Cangallo mediante imágenes Sentinel-2A. Bajo un enfoque cuantitativo y un diseño experimental categorizado como cuasi-experimento tecnológico, se procesó un dataset de 1,327 parches geoespaciales de 256×256 píxeles. Para el proceso de aprendizaje se emplearon 1,115 parches (con un 15% destinado a validación), reservando un conjunto ciego de 212 parches para la fase de prueba, lo que consolidó una muestra neta de 13,893,632 píxeles. El algoritmo fue regularizado mediante técnicas de Data Augmentation, funciones de activación LeakyReLU y una función de pérdida híbrida. En la evaluación final, el modelo alcanzó una exactitud global del 94.64% y un Índice de Intersección sobre Unión (IoU) de 0.8739. La prueba de Chi-Cuadrado de Pearson ($p < 0.05$; $V = 0.9356$) demostró una asociación estadística significativa entre las predicciones y la verdad terreno, mientras que la prueba de McNemar ($p < 0.05$) confirmó la diferencia de rendimiento frente al índice MNDWI al registrar 489,834 píxeles discrepantes a favor de la red neuronal. Como limitación del modelo implementado, se generaron 53,329 falsos positivos originados por la interferencia radiométrica de las sombras orográficas andinas. Se concluye que la arquitectura U-Net es una herramienta eficaz para la cartografía hídrica automatizada en regiones altoandinas.

Palabras clave: Cangallo, Cuerpos de agua, Deep Learning, Diseño experimental, Limitación del modelo, U-Net.

ABSTRACT

This research evaluated a Deep Learning model based on the U-Net architecture for the automated segmentation of water bodies in the province of Cangallo using Sentinel-2A images. Under a quantitative approach and an experimental design classified as a technological quasi-experiment, a dataset of 1,327 geospatial patches of 256×256 pixels was processed. For the learning process, 1,115 patches were used (with 15% allocated for validation), while a blind set of 212 patches was reserved for the testing phase, consolidating a net sample of 13,893,632 pixels. The algorithm was regularized using Data Augmentation techniques, LeakyReLU activation functions, and a hybrid loss function. In the final evaluation, the model achieved an overall accuracy of 94.64% and an Intersection over Union (IoU) Index of 0.8739. Pearson's Chi-Square test ($p < 0.05$; $V = 0.9356$) demonstrated a significant statistical association between predictions and ground truth, while the McNemar test ($p < 0.05$) confirmed the performance difference against the MNDWI index, registering 489,834 discrepant pixels in favor of the neural network. As a limitation of the implemented model, 53,329 false positives were generated due to the radiometric interference of Andean orographic shadows. It is concluded that the U-Net architecture is an effective tool for automated water mapping in high Andean regions.

Keywords: Cangallo, Deep Learning, Experimental design, Model limitation, U-Net, Water bodies.

INTRODUCCIÓN

El monitoreo de los recursos hídricos en los ecosistemas altoandinos es una tarea prioritaria frente a la variabilidad climática actual. Si bien la teledetección satelital permite la observación continua del territorio, su aplicación en zonas de topografía accidentada como la provincia de Cangallo enfrenta limitaciones técnicas significativas. Los métodos tradicionales de detección, basados en índices espectrales como el NDWI y MNDWI, suelen generar falsos positivos al confundir sombras orográficas con cuerpos de agua, o fallan al delimitar bordes irregulares, lo que demanda costosas correcciones manuales.

Ante esta problemática, la Inteligencia Artificial ofrece una solución prometedora mediante el Aprendizaje Profundo (Deep Learning). Específicamente, las Redes Neuronales Convolucionales (CNN) con arquitectura U-Net han demostrado ser un estándar sobresaliente para la segmentación semántica, ya que son capaces de aprender patrones espaciales complejos y discriminar características morfológicas más allá de la simple respuesta espectral.

La presente investigación tiene como objetivo principal determinar e implementar un modelo U-Net para automatizar la detección precisa de cuerpos de agua en la provincia de Cangallo. El estudio integra un flujo de trabajo metodológico riguroso que abarca desde la adquisición y el preprocesamiento de imágenes multiespectrales (Sentinel-2A) mediante técnicas de aumento de datos (Data Augmentation), hasta el entrenamiento en entornos computacionales en la nube de alto rendimiento. El aporte central de esta tesis radica en la obtención de un modelo robusto que alcanzó un Índice de Intersección sobre Unión (IoU) superior al 87.39%, cuya superioridad operativa y deductiva frente a los métodos convencionales fue validada estadísticamente a través de las pruebas de Chi-Cuadrado y McNemar.

DEDICATORIA

Dedico este trabajo a mis padres por su apoyo brindado durante todos estos años y por ser parte fundamental de mi formación académica y profesional.

Dedico este trabajo a mi pareja Rocio y a mi querida hija Eirene, son el motor y motivo para seguir superándome día a día.

Dedico este trabajo a mí mismo, por tener la fortaleza y la perseverancia de superarme constantemente pese a las adversidades.

AGRADECIMIENTOS

Deseo expresar mi sincera gratitud a mis padres Eugenia e Hilarión por todos sus sacrificios en beneficio de mi persona y de mis hermanos, gracias a ustedes es posible mi formación personal y profesional.

Agradezco a mi pareja Rocio por su constante apoyo, motivación y comprensión brindados durante este proceso.

Agradezco al Msc. Ing. Edward León Palacios por su guía, paciencia y conocimientos brindados. Su orientación y exigencia han sido claves para la culminación de este proyecto.

A la UNSCH y a los docentes de la E.P de Ingeniería Civil por brindarme conocimientos, estadía, confort, experiencias y una exigente formación profesional.

UNIVERSIDAD NACIONAL SAN CRISTOBAL DE HUAMANGA

AYACUCHO, ENERO DE 2026

BACH. ALEX ELVEN CHAVEZ TITO

ÍNDICE GENERAL

RESUMEN	2
ABSTRACT.....	3
INTRODUCCIÓN	4
DEDICATORIA	5
AGRADECIMIENTOS	6
ÍNDICE GENERAL	7
ÍNDICE DE FIGURAS	11
GLOSARIO	14
LISTA DE ACRÓNIMOS.....	17
LISTA DE SÍMBOLOS.....	19
CAPÍTULO I. PLANTEAMIENTO DEL PROBLEMA.....	21
1.1. Descripción del problema (mundial, país, región).....	21
1.1.1. Problemática mundial	21
1.1.2. Problemática Nacional.....	22
1.1.3. Problemática Regional.....	23
1.2. Delimitación del problema.....	24
1.2.1. Espacial.....	24
1.2.2. Temporal.....	25
1.2.3. Temática y unidad de análisis.....	25
1.3. Formulación del problema.....	25
1.3.1. Problema general.....	25
1.3.2. Problemas específicos.....	25

1.4. Justificación e importancia.	25
1.5. Limitaciones de la investigación.....	26
1.5.1. Objetivo general.....	27
1.5.2. Objetivos específicos.	27
CAPÍTULO II. MARCO TEÓRICO.	28
2.1. Antecedentes.....	28
2.1.1. Investigaciones internacionales.	28
2.1.2. Investigaciones nacionales.....	30
2.2. Bases teóricas.....	31
2.2.1. Teledetección y sensores remotos.....	31
2.2.2. Imágenes satelitales multiespectrales	32
2.2.3. Cuerpos de agua.....	35
2.2.4. Aprendizaje profundo (Deep Learning).....	39
2.3. Marco conceptual.....	47
2.3.1. Conceptos de teledetección y análisis espectral	47
2.3.2. Arquitectura y componentes del modelo Deep Learning (U-Net).....	48
2.3.3. Realización, entrenamiento y optimización.....	49
2.3.4. Evaluación y métricas de desempeño	50
CAPÍTULO III. MÉTODO DE LA INVESTIGACIÓN.....	51
3.1. Enfoque.....	51
3.2. Alcance.	51
3.3. Diseño de investigación.	51
3.4. Población y muestra.....	52

3.4.1. Población.	52
3.4.2. Muestra.	52
3.5. Hipótesis.	54
3.5.1. Hipótesis general.....	54
3.5.2. Hipótesis específicas.....	54
3.6. Operacionalización de variables.	54
3.6.1. Definición conceptual y operacional.	54
3.6.2. Operacionalización de variable independiente y dependiente	55
3.7. Técnicas e instrumentos.....	56
3.7.1. Técnicas.	56
3.7.2. Instrumentos.....	57
3.8. Técnicas estadísticas para el procesamiento de la información.....	59
3.8.1. Métricas de la evaluación del desempeño (segmentación).....	59
3.8.2. Índice de Jaccard (IoU).....	60
3.8.3. Coeficiente de coincidencia de Dice.....	60
3.8.4. Estadística inferencial y prueba de hipótesis	60
3.8.5. Herramientas de procesamiento estadístico	64
3.9. Desarrollo del trabajo de tesis.....	65
3.9.1. Fase I: Adquisición y preprocesamiento de datos geoespaciales.	65
3.9.2. Fase II: Diseño y configuración de la arquitectura U-Net.	72
3.9.3. Fase III: Entrenamiento y optimización.....	78
3.9.4. Fase IV: Inferencia y evaluación con el conjunto de prueba	82
CAPÍTULO IV. RESULTADOS.	84

4.1. Contrastación de hipótesis.	84
4.1.1. Contrastación de la hipótesis general.....	84
4.1.2. Contrastación de hipótesis específica 1	86
4.1.3. Contrastación de hipótesis específica 2	87
4.1.4. Contrastación de hipótesis específica 3	87
4.2. Análisis e interpretación de resultados.	88
4.2.1. Análisis del proceso de entrenamiento (estabilidad y regularización).	88
4.2.2. Evaluación cuantitativa y el impacto del preprocesamiento.....	91
4.2.3. Superioridad frente a métodos tradicionales (U-Net vs MNWI).....	94
4.2.4. Evaluación cualitativa (inspección visual)	95
CAPÍTULO V. CONCLUSIONES.	102
5.1. Conclusiones.....	102
5.2. Recomendaciones.	103
5.3. Trabajos futuros.	105
REFERENCIAS BIBLIOGRÁFICAS.	107

ÍNDICE DE FIGURAS

Figura 1. Ubicación de la investigación	24
Figura 2. Teledetección mediante sensores remotos	32
Figura 3. Bandas generadas por el sensor OLI.....	34
Figura 4. Imágen satelital de un lago tomada en diferentes épocas.....	36
Figura 5. Índices espectrales de la luz	38
Figura 6. Deep Learning y Machine Learning dentro de la Inteligencia Artificial.	39
Figura 7. Funcionamiento de una red neuronal convolucional (CNN).	40
Figura 8. Arquitectura típica de un modelo U-Net.....	46
Figura 9. Descarga de granula de Sentinel-2A des de Copernicus Browser.	66
Figura 10. Imagen multiespectral obtenida por la composición de bandas en ArcGIS Pro	68
Figura 11. Mascara binaria de la imagen satelital multiespectral obtenida en ArcGis Pro	71
Figura 12. Implementación de la capa de entrada del modelo U-Net	73
Figura 13. Implementación del Encoder, Bottlenet y Decoder del modelo U-Net.....	76
Figura 14. Representación Topológica de la arquitectura del modelo U-Net implementado.....	76
Figura 15. Entrenamiento del modelo de segmentación semántica U-Net.....	81
Figura 16. Inferencia visual del modelo en el test de prueba	83
Figura 17. Curva de aprendizaje y curva de precisión del modelo entrenado con aumentación de datos.....	89
Figura 18. Curva de aprendizaje y curva de precisión del modelo entrenado sin aumentación de datos.....	90
Figura 19. Matriz de confusión del modelo entrenado con aumentación de datos.....	92

Figura 20. Matriz de confusión del modelo entrenado sin aumentación de datos.....	92
Figura 21. Predicciones del modelo U-Net vs MNDWI.....	95

ÍNDICE DE TABLAS

Tabla 1. Bandas generadas por el sensor OLI.....	34
Tabla 2. Bandas generadas por el sensor MSI.	35
Tabla 3. Definiciones de las variables	55
Tabla 4. Operacionalización de la Variable Independiente.	55
Tabla 5. Operacionalización de la Variable Dependiente.....	56
Tabla 6. Comparativa de estabilidad durante el entrenamiento (Momentos críticos). ..	90
Tabla 7. Impacto del preprocesamiento en el rendimiento del conjunto de prueba (Test Set).....	93
Tabla 8. Tabla de contingencia de la prueba de McNemar (U-Net vs. MNDWI).....	94

GLOSARIO

Aprendizaje Profundo (Deep Learning): Subdisciplina del aprendizaje automático (Machine Learning) basado en redes neuronales artificiales con múltiples capas ocultas, capaces de aprender representaciones jerárquicas de los datos con altos niveles de abstracción.

Aumento de Datos (Data Augmentation): Técnica de regularización utilizada en el entrenamiento de modelos de aprendizaje profundo que consiste en generar nuevas muestras sintéticas a partir de los datos originales mediante transformaciones geométricas (rotación, traslación, volteo), con el fin de aumentar la diversidad del conjunto de entrenamiento y reducir el sobreajuste.

Capas Convolucionales: Componentes fundamentales de las redes neuronales convolucionales (CNN) que aplican filtros matemáticos a la imagen de entrada para extraer características espaciales como bordes, texturas y formas.

Coefficiente Dice: Métrica estadística utilizada para medir la similitud entre dos conjuntos de datos. En segmentación de imágenes, cuantifica la superposición espacial entre la máscara predicha y la máscara real; es matemáticamente equivalente al F1-Score.

Entropía Cruzada Binaria (Binary Cross-entropy): Función de pérdida utilizada en problemas de clasificación binaria. Mide la discrepancia entre la distribución de probabilidad predicha por el modelo y la distribución real (etiqueta verdadera), penalizando las predicciones incorrectas con alta confianza.

Época (Epoch): Hiperparámetro que define un ciclo completo en el que todo el conjunto de datos de entrenamiento pasa hacia adelante y hacia atrás (retropropagación) a través de la red neuronal una sola vez.

Función de Activación: Función matemática aplicada a la salida de una neurona artificial para introducir no linealidad en el modelo, permitiéndole aprender patrones complejos. Ejemplos comunes son ReLU (Rectified Linear Unit) y Sigmoid.

Hiperparámetros: Variables de configuración externa al modelo cuyo valor se establece antes de iniciar el proceso de entrenamiento (ej. tasa de aprendizaje, tamaño de lote, número de filtros) y que no se aprenden directamente de los datos.

Imagen Satelital: Representación visual de la superficie terrestre capturada por sensores montados en satélites artificiales, que registran la radiación electromagnética en diferentes bandas del espectro.

Pixel: Unidad básica de una imagen digitalizada en pantalla a base de puntos de color o en escala de grises.

Normalización: Proceso de ajustar los datos para que tengan una escala común o estandarizada, lo que facilita su comparación y análisis.

Intersección sobre Unión (IoU): También conocido como Índice de Jaccard. Es la métrica estándar para evaluar la precisión en la segmentación semántica. Se calcula dividiendo el área de la intersección entre la predicción y la verdad terreno por el área de su unión.

Máscara Binaria: Imagen en la que cada píxel tiene solo dos valores posibles (generalmente 0 y 1 o blanco y negro), utilizada para etiquetar las áreas de interés (ej. 1=Agua) frente al fondo (ej. 0=No Agua).

Negative Mining: Estrategia de preprocesamiento o entrenamiento que consiste en seleccionar o filtrar ejemplos negativos (fondo) difíciles o irrelevantes para mejorar la capacidad de discriminación del modelo.

Retropropagación (Backpropagation): Algoritmo fundamental para el entrenamiento de redes neuronales que calcula el gradiente de la función de pérdida respecto a los pesos de la red, permitiendo ajustarlos iterativamente para minimizar el error.

Segmentación Semántica: Tarea de visión por computadora que consiste en clasificar cada píxel de una imagen en una categoría específica, permitiendo delimitar con precisión los contornos de los objetos de interés.

Sobreajuste (Overfitting): Fenómeno de error de modelado que ocurre cuando una red neuronal aprende "de memoria" los detalles y el ruido de los datos de entrenamiento en lugar de los patrones generales, perdiendo capacidad para generalizar frente a datos nuevos.

Verdad Terreno (Ground Truth): Información recolectada mediante observación directa o validación experta que sirve como referencia absoluta o "etiqueta correcta" para entrenar y evaluar algoritmos de aprendizaje supervisado.

LISTA DE ACRÓNIMOS

ANA	: Autoridad Nacional del Agua.
API	: Application Programming Interface (Interfaz de Programación de Aplicaciones).
CNN	: Convolutional Neural Network (Red Neuronal Convolutacional).
DL	: Deep Learning (Aprendizaje Profundo).
FN	: Falsos Negativos (False Negatives).
FP	: Falsos Positivos (False Positives).
GIS	: Geographic Information System (Sistema de Información Geográfica).
GPU	: Graphics Processing Unit (Unidad de Procesamiento Gráfico).
IA	: Inteligencia Artificial.
IGN	: Instituto Geográfico Nacional.
IoU	: Intersection over Union (Intersección sobre Unión).
NDWI	: Normalized Difference Water Index (Índice Diferencial de Agua Normalizado).
MNDWI	: Modified Normalized Difference Water Index (Índice Diferencial de Agua Normalizado Modificado).
NASA	: National Aeronautics and Space Administration (Administración Nacional de Aeronáutica y el Espacio).
NIR	: Near Infrared (Infrarrojo Cercano).
RAM	: Random Access Memory (Memoria de Acceso Aleatorio).
ReLU	: Rectified Linear Unit (Unidad Lineal Rectificada).
RGB	: Red, Green, Blue (Rojo, Verde, Azul - Modelo de color).
RMSE	: Root Mean Square Error (Raíz del Error Cuadrático Medio).
RNA	: Red Neuronal Artificial.

SIG : Sistema de Información Geográfica.

SWIR : Short-Wave Infrared (Infrarrojo de Onda Corta).

TIFF : Tagged Image File Format (Formato de Archivo de Imagen Etiquetado).

TN : Verdaderos Negativos (True Negatives).

TP : Verdaderos Positivos (True Positives).

USGS : United States Geological Survey (Servicio Geológico de los Estados Unidos).

UTM : Universal Transverse Mercator (Sistema de Coordenadas Universal Transversal de Mercator).

LISTA DE SÍMBOLOS

Acc	: Exactitud global del modelo (<i>Accuracy</i>).
α	: Tasa de aprendizaje lo nivel de significancia.
β_1, β_2	: Hiperparámetros de decaimiento exponencial del optimizador Adam.
χ^2	: Estadístico de la prueba de independencia Chi-Cuadrado de Pearson.
$Dice$	: Coeficiente de similitud Dice.
ϵ	: Epsilon, valor infinitesimal (por ejemplo $1e^{-7}$).
$\in$	: Símbolo de pertenencia a un conjunto
H_0	: Hipótesis nula en las pruebas estadísticas.
H_1	: Hipótesis alterna o del investigador.
IoU	: Índice de Intersección sobre Unión (<i>Intersection over Union</i>).
$\mathcal{L}$	: Función de costo o pérdida (<i>Loss Function</i>).
N	: Tamaño total de la muestra o número de imágenes en el dataset.
p	: Valor que determina la significancia estadística de una prueba.
R	: Conjunto de los números reales.
$\sigma(z)$	: Función de activación Sigmoide.
τ	: Umbral de decisión (<i>Threshold</i>).
∇	: Operador gradiente.
x	: Variable de entrada (tensor de la imagen satelital).
y	: Variable objetivo o etiqueta real (<i>Ground Truth</i>).
$\hat{y}$	: Predicción generada por la red neuronal.
w	: Matriz de pesos sinápticos (<i>weights</i>) de las capas convolucionales.

H	: altura de imagen de entrada en pixeles.
W	: Ancho de imagen de entrada en pixeles.
C	: Numero de canales de la imagen.
ρ_{λ}	: Reflectancia en una longitud de onda especifica.

CAPÍTULO I. PLANTEAMIENTO DEL PROBLEMA

1.1. Descripción del problema (mundial, país, región).

1.1.1. Problemática mundial

A nivel global, la seguridad y la gestión sostenible de los recursos hídricos superficiales se enfrentan a una crisis sin precedentes históricos. El United Nations World Water Assessment Programme [WWAP] (2023) advierte que el acceso desigual a los servicios hídricos esenciales está precarizando la salud, los medios de subsistencia y el desarrollo equitativo de las sociedades. Esta situación se agrava de manera acelerada debido al crecimiento demográfico y a los desequilibrios en el uso de la tierra. El impacto es de tal magnitud que el cumplimiento del Objetivo de Desarrollo Sostenible 6 (SDG 6), orientado a garantizar agua limpia y saneamiento para todos hacia el año 2030, se encuentra bajo un severo compromiso estructural, lo que exige herramientas de monitorización ambiental de escala masiva y alta frecuencia temporal.

Esta crisis en el suministro y preservación del recurso vital se ve directamente intensificada por los efectos antropogénicos en el clima global. El Intergovernmental Panel on Climate Change [IPCC] (2023) señala con un alto nivel de confianza que las alteraciones térmicas globales están desestabilizando críticamente el ciclo hidrológico mundial, acelerando la pérdida de la criosfera, extendiendo los periodos de sequía extrema e incrementando la destructividad de las inundaciones repentinas. Alrededor de la mitad de la población del planeta experimenta escasez grave de agua durante diversas épocas del año, lo que convierte la delimitación geográfica y temporal de los cuerpos de agua superficiales en una tarea prioritaria para la prevención de desastres y la planificación de la resiliencia comunitaria. Sin embargo, la velocidad con la que estas masas de agua dulce cambian su morfología supera la capacidad operativa de las metodologías de inventario convencionales.

Históricamente, los sistemas de observación de la Tierra y la teledetección espacial han representado la solución óptima para estudiar estas dinámicas a gran escala. El trabajo cartográfico global desarrollado por Pekel et al. (2016) marcó un hito en la disciplina al procesar millones de imágenes satelitales durante tres décadas, demostrando el valor de los sensores ópticos para registrar las variaciones espaciotemporales del agua superficial. No obstante, las aproximaciones clásicas basadas en álgebra de bandas e índices espectrales con umbrales fijos adolecen de severas limitaciones técnicas. Estos algoritmos empíricos suelen fallar o inducir a errores de sobreestimación ante la presencia de nubes, aerosoles atmosféricos, cuerpos hídricos de alta turbidez o sombras densas proyectadas por el relieve topográfico y la infraestructura urbana.

Con el fin de superar estos cuellos de botella metodológicos, la comunidad científica internacional ha impulsado una transición hacia técnicas automatizadas avanzadas de análisis digital. Zhu et al. (2017) exponen que el procesamiento inteligente de datos geoespaciales mediante algoritmos de aprendizaje profundo (Deep Learning) permite extraer patrones semánticos y morfológicos abstractos de forma totalmente autónoma, superando la rigidez de la clasificación basada en píxeles aislados

1.1.2. Problemática Nacional

A nivel nacional, el Perú presenta una marcada asimetría espacial en la disponibilidad de sus recursos hídricos; aunque la vertiente del Pacífico alberga a más del 60 % de la población y concentra la principal actividad agroexportadora, dispone de apenas el 1.8 % de la oferta hídrica total, lo que condiciona al territorio a una alta vulnerabilidad frente a sequías sostenidas (Autoridad Nacional del Agua [ANA], 2021). Para mitigar este déficit estacional, las lagunas altoandinas actúan como los principales reservorios naturales que regulan el ciclo hidrológico de las cuencas; no obstante, el Instituto Nacional de Investigación en Glaciares y Ecosistemas de Montaña (INAIGEM, 2018) advierte que el

monitoreo y la actualización del Inventario Nacional de Lagunas enfrentan barreras logísticas y técnicas críticas debido a la agreste topografía de la Cordillera de los Andes. Históricamente, la cartografía hídrica estatal ha dependido de inspecciones in situ y métodos de teledetección convencionales que resultan ineficaces en la geografía andina, ya que los algoritmos de índices espectrales clásicos son incapaces de discriminar las densas sombras orográficas proyectadas por los relieves escarpados, generando mapas con altos porcentajes de falsos positivos y subestimando las reservas reales en épocas de estiaje (Ministerio del Ambiente [MINAM], 2022). En consecuencia, el país enfrenta una urgente brecha metodológica: es imperativo que las instituciones peruanas responsables de la gestión hídrica transicionen hacia la adopción de modelos de inteligencia artificial y aprendizaje profundo que logren aislar el ruido topográfico, garantizando la automatización de un mapeo satelital de alta fidelidad que respalde la seguridad hídrica nacional.

1.1.3. Problemática Regional

A nivel regional, el departamento de Ayacucho representa uno de los territorios altoandinos más expuestos a la variabilidad climática extrema, enfrentando episodios recurrentes de déficit hídrico severo que amenazan la sostenibilidad de sus proyectos de irrigación y la seguridad alimentaria de sus comunidades (Centro Nacional de Estimación, Prevención y Reducción del Riesgo de Desastres [CENEPRED], 2023). Dentro de esta jurisdicción, a nivel local, la provincia de Cangallo constituye un escenario de riesgo crítico; la agudización de sus periodos de estiaje y la falta de precipitaciones han requerido la intervención directa del gobierno central mediante reiteradas declaratorias de Estado de Emergencia (Presidencia del Consejo de ministros [PCM], 2023). Para mitigar este impacto y formular proyectos de ingeniería hidráulica eficientes, las autoridades locales necesitan monitorear con urgencia sus cuerpos de agua

superficiales, tales como las lagunas altoandinas y "ccochas". Sin embargo, la geografía altamente escarpada de Cangallo imposibilita la actualización constante de inventarios mediante levantamientos topográficos de campo. Sumado a ello, la aplicación de herramientas geomáticas tradicionales en esta provincia resulta inoperante: los algoritmos basados en índices espectrales convencionales fracasan estrepitosamente debido a que confunden sistemáticamente las profundas sombras orográficas proyectadas por los relieves de la zona con espejos de agua, generando una cartografía local inservible para la toma de decisiones (Gobierno Regional de Ayacucho [GORE Ayacucho], 2022). En consecuencia, focalizar la investigación en la provincia de Cangallo se justifica por la necesidad técnica imperativa de implementar un modelo de aprendizaje profundo (deep learning) que logre superar el ruido orográfico local.

1.2. Delimitación del problema.

1.2.1. Espacial.

La presente investigación se desarrollará en la provincia de Cangallo.

Región: Ayacucho.

Provincia: Cangallo.

Figura 1. Ubicación de la investigación



Nota: elaboración propia

1.2.2. Temporal.

La investigación se desarrollará en el periodo del año 2024.

1.2.3. Temática y unidad de análisis.

la detección de cuerpos de agua en imágenes satelitales multiespectrales.

1.3. Formulación del problema.

1.3.1. Problema general.

¿Cómo influye la implementación de un modelo de Deep Learning en la precisión de la detección automatizada de cuerpos de agua mediante el procesamiento de imágenes satelitales en la provincia de Cangallo, 2024?

1.3.2. Problemas específicos.

1. ¿Cuál es la precisión del modelo U-Net en la detección de cuerpos de agua mediante imágenes satelitales?
2. ¿Qué nivel de desempeño presenta el modelo U-Net en la segmentación de cuerpos de agua según métricas IoU (Intersection over Union) y Accuracy (exactitud)?
3. ¿Existe una diferencia significativa entre la detección de cuerpos de agua mediante el modelo U-Net y el índice espectral MNDWI (Modified Normalized Difference Water Index)?

1.4. Justificación e importancia.

La presente investigación se fundamenta en la necesidad imperativa de modernizar los procesos de monitoreo de recursos hídricos en entornos geográficamente complejos. La relevancia del estudio se sustenta en las siguientes dimensiones:

Justificación técnica y tecnológica: La metodología propuesta trasciende las limitaciones de la teledetección tradicional basada en índices de umbral fijo, como el NDWI, los cuales suelen presentar errores de clasificación en zonas de relieve accidentado debido a las sombras proyectadas por la topografía andina.

Justificación científica: El estudio contribuye al estado del arte en la aplicación de modelos de Deep Learning para la hidrología y la ingeniería civil en el Perú. Al validar el desempeño de redes neuronales convolucionales profundas en la provincia de Cangallo, se genera un precedente científico sobre la eficacia de estas arquitecturas en climas de alta montaña y topografías accidentadas.

Prevención de riesgos: la detección precisa de masas de agua es fundamental para el manejo y prevención de desastres. El modelo desarrollado facilita la identificación de áreas vulnerables a inundaciones o procesos de sequía extrema, permitiendo una respuesta rápida y efectiva ante eventos climáticos adversos.

Justificación ambiental y de sostenibilidad: En el marco del cambio climático, la preservación de los ecosistemas acuáticos es crítica. Esta investigación aporta a la conservación ambiental al permitir un seguimiento detallado de las variaciones en los espejos de agua de la provincia de Cangallo, contribuyendo a la protección de la biodiversidad y al mantenimiento del equilibrio hidrológico local.

1.5. Limitaciones de la investigación.

El desarrollo del presente estudio se ve influenciado por diversos factores técnicos, ambientales y operativos que delimitan el alcance de los resultados obtenidos.

- El acceso a datos, la disponibilidad y la calidad de las imágenes satelitales pueden variar, lo que puede afectar el modelo.

- La capacidad computacional, el procesamiento de grandes volúmenes de datos requiere infraestructura tecnológica avanzada y recursos computacionales significativos.
- Los factores climáticos, la variabilidad climática y las condiciones meteorológicas pueden influir en la calidad de las imágenes satelitales y en la precisión de los cuerpos de agua.

1.5.1. Objetivo general.

Evaluar la influencia de la implementación de un modelo de Deep Learning en la precisión de la detección automatizada de cuerpos de agua mediante el procesamiento de imágenes satelitales en la provincia de Cangallo, 2024.

1.5.2. Objetivos específicos.

1. Determinar la precisión del modelo U-Net en la detección de Lagunas.
2. Evaluar el desempeño del modelo U-Net en la segmentación de Lagunas mediante métricas IoU y Accuracy.
3. Comparar estadísticamente la precisión del modelo U-Net frente al índice MNDWI mediante la prueba de McNemar.

CAPÍTULO II. MARCO TEÓRICO.

2.1. Antecedentes.

2.1.1. Investigaciones internacionales.

El campo de la teledetección hídrica ha experimentado una transición paradigmática desde los índices multispectrales hacia modelos de segmentación semántica profunda.

A continuación, se detallan los pilares científicos internacionales para esta tesis:

Isikdogan et al. (2017), en su influyente investigación "Surface water mapping by deep learning", abordaron el desafío de cartografiar cuerpos de agua a gran escala superando las severas interferencias generadas por las sombras del terreno, la nieve y las nubes. El estudio demostró que los índices espectrales tradicionales, basados en umbrales estáticos, fallan consistentemente en topografías accidentadas. Para resolverlo, los autores desarrollaron y entrenaron una red neuronal convolucional totalmente conectada utilizando un dataset masivo de imágenes satelitales multispectrales. La investigación comprobó empíricamente que el modelo de aprendizaje profundo fue capaz de aprender las características espaciales y contextuales del agua, aislando la firma hídrica de las sombras topográficas profundas con una eficacia significativamente mayor a la de los métodos convencionales de teledetección. Este antecedente es fundamental para la presente tesis, ya que fundamenta metodológicamente que el Deep Learning es la tecnología idónea para discriminar recursos hídricos en zonas donde el relieve agreste proyecta sombras orográficas que confunden a los algoritmos clásicos.

Por su parte, Feng et al. (2019) desarrollaron el artículo "Water body extraction from very high-resolution remote sensing imagery using deep U-Net and a superpixel-based conditional random field model". Esta investigación se centró específicamente en evaluar el rendimiento de la arquitectura U-Net frente a la dificultad técnica de delinear fronteras hídricas con alta precisión. Los autores entrenaron la red U-Net procesando imágenes

satelitales de alta resolución y demostraron que las conexiones saltadas (skip connections) de esta arquitectura son vitales para recuperar los bordes espaciales finos que se pierden durante la extracción de características. Al combinar las predicciones de la U-Net con modelos de posprocesamiento, lograron segmentar el agua de manera excepcionalmente nítida, superando métricas de precisión del 90 %. Este estudio justifica de manera directa la elección de la red U-Net para la presente investigación, validándola como la arquitectura convolucional más eficiente para delimitar geometrías irregulares típicas de los reservorios naturales.

De manera complementaria, Chen et al. (2018), en su investigación "Extraction of urban water bodies from high-resolution remote-sensing imagery using deep learning", evaluaron la capacidad de las redes neuronales convolucionales para extraer masas de agua en entornos con alto ruido visual y falsos positivos, comparando su rendimiento contra el Índice Diferencial de Agua Normalizado (NDWI). El estudio se ejecutó procesando bandas multiespectrales y demostró que, mientras el NDWI confundía constantemente las oscuras sombras de las infraestructuras con agua, la red convolucional logró abstraer el contexto espacial del píxel, mitigando drásticamente el error de comisión. Este trabajo aporta un sustento técnico crucial al presente proyecto, demostrando que para detectar cuerpos de agua en escenarios con interferencia radiométrica como el ruido topográfico de la sierra peruana es necesario abandonar las técnicas basadas únicamente en umbrales de píxeles (índices) y transicionar hacia modelos que evalúen el contexto del terreno.

Finalmente, en el ámbito latinoamericano, las técnicas de teledetección espacial han sido fundamentales para evaluar los impactos del cambio climático y la escasez hídrica extrema. Mancilla Fuentes (2021) buscó detectar y cuantificar los drásticos cambios superficiales ocurridos en la Laguna de Aculeo y el Lago Peñuelas, en Chile, durante la

década de 2010 a 2020. Para cumplir este propósito, el autor implementó el Análisis de Componentes Principales (ACP) aplicado a datos espaciales y evaluó el desempeño de los índices espectrales NDWI y MNDWI extraídos de imágenes multitemporales Landsat. El estudio comprobó una alarmante tendencia a la disminución de la superficie de agua debido a la prolongada megasequía que azota la región, confirmando que la Laguna de Aculeo se secó por completo en 2018 y que el Lago Peñuelas sufrió una pérdida progresiva que superó el 60% de su espejo de agua original, llegando incluso a más del 80% hacia 2020. Además, desde el punto de vista metodológico, el autor concluyó que el índice NDWI, en conjunto con el método estadístico ACP, demostraron ser las herramientas más precisas para detectar las variaciones espaciotemporales en este tipo de cuerpos de agua.

2.1.2. Investigaciones nacionales.

En el Perú, la aplicación de Deep Learning para recursos hídricos es casi nula. El 95 % de las tesis peruanas de ingeniería civil o geografía miden las lagunas usando métodos tradicionales (MNDWI, NDWI procesados en ArcGIS o QGIS).

Ramos Barzola (2021), en un estudio enfocado en la Cordillera Blanca en Perú, utilizó técnicas de teledetección para estimar la pérdida de cobertura glaciar en la cuenca del río Santa. La metodología incluyó el uso de algoritmos de Random Forest e índices de nieve y temperatura de superficie aplicados a productos Landsat y Sentinel-2. Los hallazgos determinaron que la clasificación supervisada era altamente precisa para predecir la pérdida de masa de hielo, proyectando escenarios hasta el año 2145. Este antecedente proporciona un contexto regional sobre la dinámica hídrica en los Andes peruanos, reforzando la necesidad de monitoreo constante en zonas de alta montaña.

García Tadeo (2023) investigó la detección de avalanchas de origen glaciar en el ambiente proglacial de Palcacocha. El estudio empleó análisis SAR (Synthetic Aperture

Radar) de la misión Sentinel-1 para identificar variaciones en glaciares como el Palcaraju y Pucaranra, encontrando correlaciones significativas entre el flujo de radón, la temperatura y la humedad. Una contribución relevante de este trabajo es la demostración de que las imágenes SAR permiten el monitoreo hídrico sin restricciones de iluminación solar, complementando los datos ópticos en zonas de alta nubosidad.

Palma Gongora (2023) analizó la evolución temporal de la contaminación por coliformes fecales en aguas costeras mediante imágenes de los sensores Landsat 7 ETM+ y Landsat 8 OLI. La investigación estableció una correlación directa entre la reflectancia de bandas espectrales específicas y los niveles de microorganismos, permitiendo generar mapas de distribución espacial en la bahía de Pucusana. Este estudio evidencia que el procesamiento digital de imágenes satelitales no solo es útil para la detección de masas de agua, sino también para inferir características cualitativas del recurso hídrico a través de modelos de predicción espacial.

2.2. Bases teóricas.

Esta sección fundamenta los principios físicos de la teledetección y los pilares matemáticos del aprendizaje profundo que permiten la detección automatizada de cuerpos de agua.

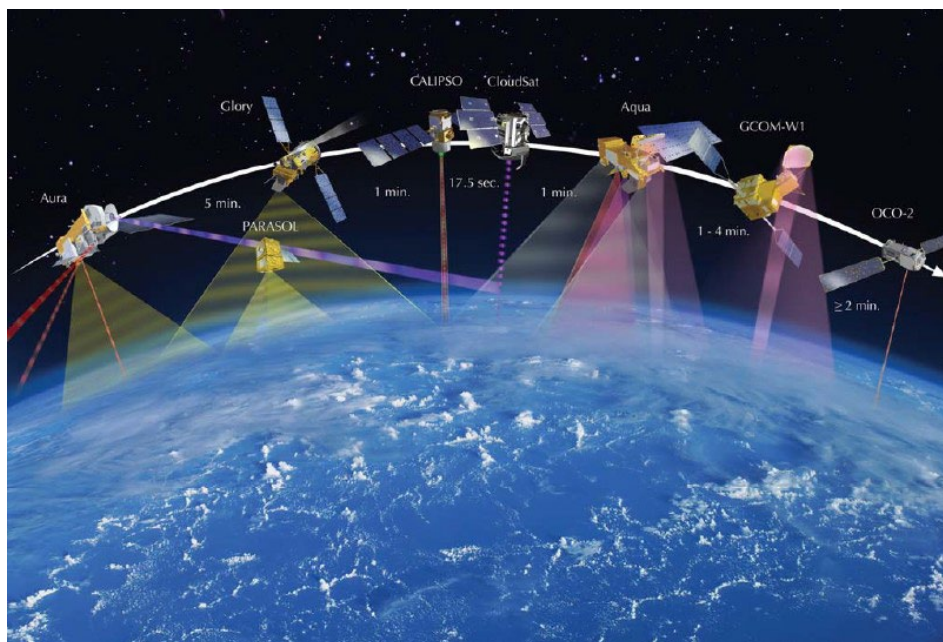
2.2.1. Teledetección y sensores remotos

La teledetección es la ciencia y tecnología de obtener información sobre objetos o áreas a distancia, utilizando sensores que detectan y miden la radiación electromagnética reflejada o emitida por los objetos. Esta técnica permite la observación y el análisis de la superficie terrestre sin necesidad de contacto físico directo. Los sensores remotos son dispositivos que se instalan en plataformas como satélites, aviones o drones, y tienen la capacidad de recolectar información sobre la superficie terrestre a través de la detección

de radiación electromagnética en diferentes longitudes de onda. Tenemos dos tipos de sensores remotos:

- **Sensores pasivos:** Detectan la radiación reflejada o emitida por los objetos. Incluyen los sensores ópticos e infrarrojos.
- **Sensores activos:** Emiten su propia energía y miden la radiación reflejada. Incluyen el radar y el LIDAR.

Figura 2. Teledetección mediante sensores remotos



Nota: captura de imágenes satelitales utilizando sensores remotos

2.2.2. Imágenes satelitales multispectrales

Las imágenes satelitales multispectrales se definen como representaciones digitales de la superficie terrestre compuestas por múltiples capas o bandas, cada una de las cuales captura datos en rangos específicos del espectro electromagnético. A diferencia de las imágenes ópticas convencionales, estas integran longitudes de onda que se encuentran fuera del espectro visible por el ojo humano, lo que permite identificar propiedades físicas y químicas de los objetos que no son perceptibles a simple vista.

2.2.2.1. Composición de bandas

Una imagen multiespectral se compone de varias capas o bandas, donde cada una representa un rango específico de longitudes de onda. En el contexto de esta investigación, se destacan las siguientes:

- **Bandas Visibles (RGB):** Comprenden el Rojo (Red), Verde (Green) y Azul (Blue). Proporcionan información sobre el color y la textura superficial.
- **Infrarrojo Cercano (NIR):** Es la banda más crítica para la detección de recursos hídricos. Debido a que el agua absorbe casi la totalidad de la energía en esta región mientras que la vegetación la refleja intensamente, el NIR permite una discriminación nítida de los límites de cuerpos de agua.
- **Infrarrojo de Onda Corta (SWIR):** Utilizada para diferenciar la humedad del suelo y tipos de rocas.

Las imágenes multiespectrales más utilizadas actualmente provienen de las misiones Landsat 8/9 (Sensor OLI y TIRS) y Sentinel-2A (sensor MSI).

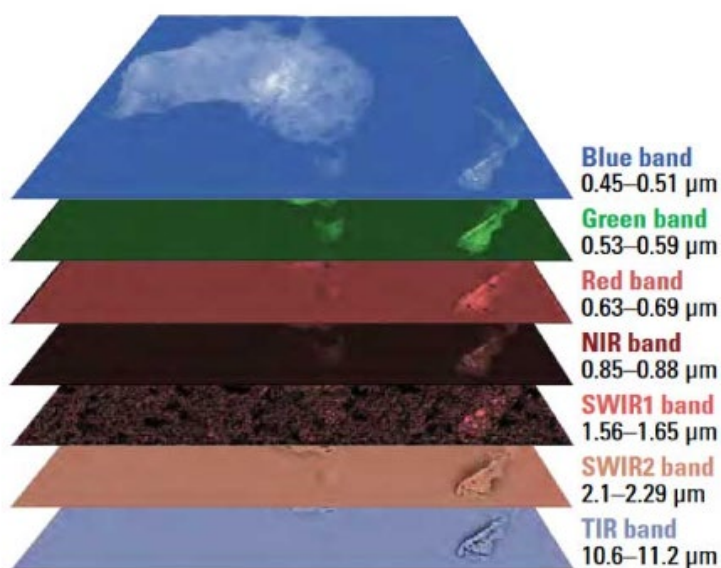
El programa Landsat, gestionado por la USGS, ofrece una resolución espacial de 30 metros para sus bandas multiespectrales. Es el estándar para estudios históricos de cuerpos de agua.

Tabla 1. Bandas generadas por el sensor OLI.

Banda	Nombre	Longitud de Onda (μm)	Resolución
B1	Aerosoles costeros	0.43 – 0.45	30 m.
B2	Azul (Blue)	0.45 – 0.51	30 m.
B3	Verde (Green)	0.53 – 0.59	30 m
B4	Rojo (Red)	0.64 – 0.67	30 m
B5	Infrarrojo Cercano (NIR)	0.85 – 0.88	30 m
B6	SWIR 1	1.57 – 1.65	30 m
B7	SWIR 2	2.11 – 2.29	30 m
B8	Pancromático	0.50 – 0.68	30 m
B9	Cirrus	1.36 – 1.39	30 m

Nota: adaptado de GISGeography

Figura 3. Bandas generadas por el sensor OLI



Nota: adaptado de GISGeography

El satélite Sentinel-2A (sensor MSI) es Operado por la Agencia Espacial Europea (ESA), este sensor es superior para la detección de cuerpos de agua pequeños en la provincia de Cangallo debido a su resolución de 10 metros en las bandas principales.

Tabla 2. Bandas generadas por el sensor MSI.

Banda	Nombre / Descripción	Longitud de onda central (nm)	Resolución espacial (m)
B1	Aerosoles costeros	443	60
B2	Azul (Blue)	490	10
B3	Verde (Green)	560	10
B4	Rojo (Red)	665	10
B5	Vegetación Red Edge 1	705	20
B6	Vegetación Red Edge 2	740	20
B7	Vegetación Red Edge 3	783	20
B8	NIR (Infrarrojo cercano)	842	10
B8A	NIR Estrecho (Narrow NIR)	865	20
B9	Vapor de agua	945	60
B10	SWIR - Cirros	1375	60
B11	SWIR 1 (Infrarrojo de onda corta)	1610	20
B12	SWIR 2 (Infrarrojo de onda corta)	2190	20

Nota: adaptado de sentinelhub

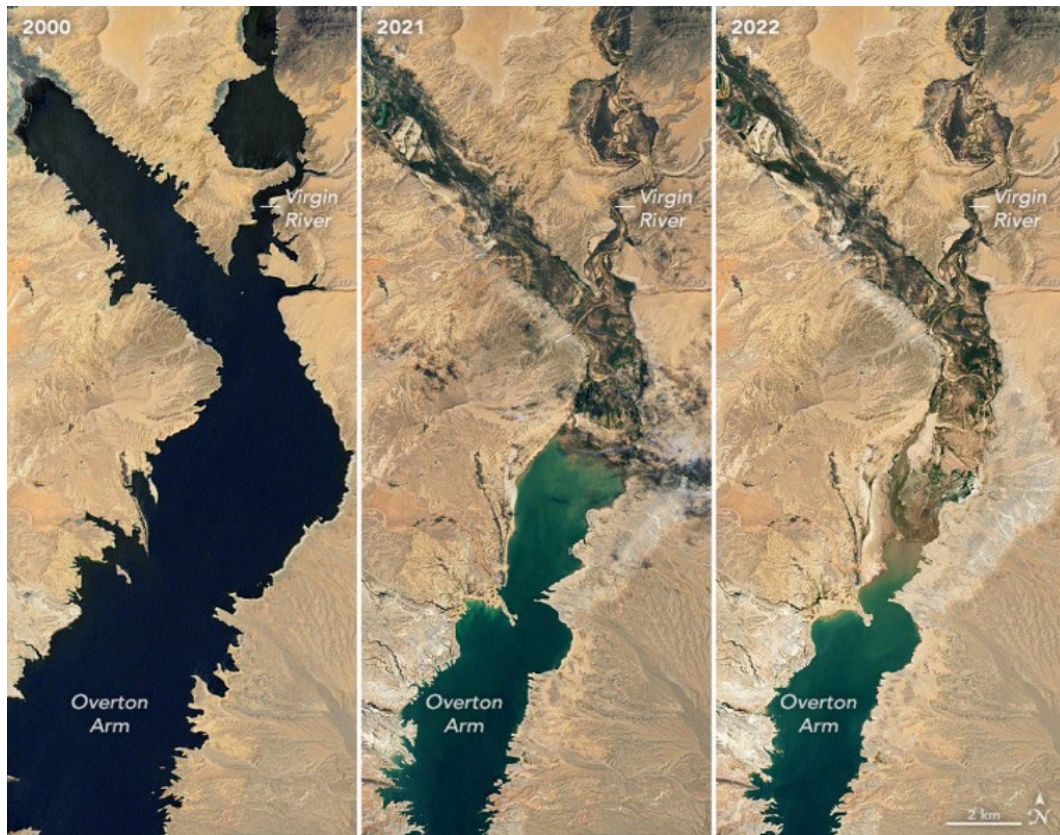
2.2.3. Cuerpos de agua

Las masas de agua son extensiones significativas de agua superficial, como lagos, ríos, embalses, océanos, mares y cuerpos de agua artificiales. Estas masas pueden ser naturales o artificiales y juegan un papel crucial en el ciclo hidrológico y en la sostenibilidad ambiental. Tenemos varios tipos de masas de agua:

- **Ríos:** Corrientes de agua que fluyen sobre la superficie terrestre.
- **Lagos:** Grandes cuerpos de agua dulce o salada, generalmente rodeados de tierra.
- **Embalses:** Cuerpos de agua artificiales creados por la construcción de presas.
- **Océanos y Mares:** Grandes extensiones de agua salada que cubren la mayor parte de la superficie terrestre.

La Investigación estará enfocada a la detección de cuerpos de agua superficiales específicamente a cuerpos de agua Lenticulares como Lagos y Lagunas.

Figura 4. Imágen satelital de un lago tomada en diferentes épocas.



Nota: obtenida de NBC NEWS

2.2.3.1. Características de los cuerpos de agua

Las características espectrales del agua se refieren a cómo el agua refleja, absorbe y transmite la radiación electromagnética en diferentes longitudes de onda. Estas características son fundamentales para la detección y el análisis de masas de agua mediante teledetección.

a) Alta absorción en el infrarrojo

Esta es la característica más importante. El agua absorbe casi toda la radiación en las regiones del infrarrojo cercano (NIR) y del infrarrojo de onda corta (SWIR).

En las imágenes de satélite que usan bandas infrarrojas, los cuerpos de agua aparecen de color negro o muy oscuro, porque apenas reflejan luz hacia el sensor.

b) Baja reflectancia general

Comparado con otros elementos de la superficie terrestre (suelo, rocas, plantas), el agua tiene una reflectancia general mucho menor. Refleja un máximo del 10% de la radiación incidente, mientras que la vegetación puede reflejar hasta un 50%.

c) Reflectancia variable en el visible

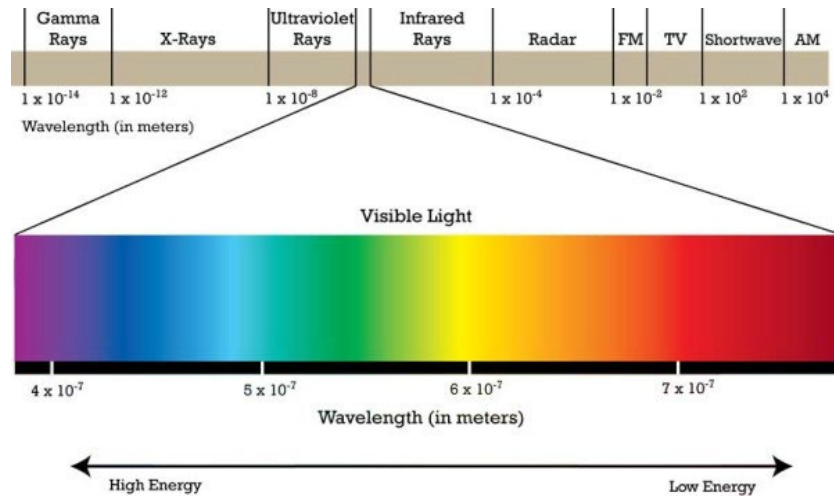
En la región visible del espectro, la reflectancia del agua es mayor, especialmente en la región azul-verde, donde la luz solar penetra más.

- **Agua Clara:** Refleja más luz azul, por lo que se ve de color azul oscuro en imágenes de color verdadero.
- **Agua Turbia con sedimentos:** Las partículas suspendidas (turbidez) dispersan y reflejan más luz en las bandas verde y rojo. El agua se ve entonces más verde, turquesa o marrón.

2.2.3.2. Índices espectrales

"Los índices espectrales son algoritmos matemáticos basados en operaciones algebraicas entre las bandas de una imagen satelital, diseñados para resaltar la firma específica de una cobertura y atenuar el ruido de otras (Jensen, 2016; Lillesand et al., 2015). En la detección hídrica, estos índices aprovechan el contraste entre la alta reflectancia del agua en el espectro visible y su absorción casi total en las regiones del infrarrojo (Chuvieco, 2020)

Figura 5. Índices espectrales de la luz



Nota: obtenida de geoespectral.com

a) Índice de Diferencia Normalizada de Agua (NDWI)

Propuesto originalmente por McFeeters, S. K. (1996), el NDWI es el estándar para resaltar masas de agua abiertas. Utiliza la banda del Verde y el Infrarrojo Cercano (NIR) para maximizar la reflectancia hídrica y minimizar la de la vegetación.

$$NDWI = \frac{(GREEN - NIR)}{(GREEN + NIR)}$$

Para Sentinel-2A se utilizan las bandas B3 (verde) y B8 (NIR).

b) Índice de Diferencia Normalizada de Agua Modificado (MNDWI)

Propuesto por Xu, H. (2006), El MNDWI sustituye la banda NIR por el Infrarrojo de Onda Corta (SWIR). Esta modificación es crucial en zonas con presencia de infraestructura, ya que el SWIR permite diferenciar mejor el agua de las áreas urbanas, las cuales suelen generar falsos positivos en el NDWI tradicional:

$$MNDWI = \frac{(GREEN - SWIR)}{(GREEN + SWIR)}$$

Para Sentinel-2A se utilizan las bandas B3 (verde) y B11 (SWIR1).

c) Índice de Extracción Automática de Agua (AWEI)

Para escenarios con sombras topográficas complejas se emplea el AWEI. Este índice combina múltiples bandas para mejorar la precisión en áreas donde la sombra de las montañas se confunde espectralmente con el agua (Feyisa et al., 2014).

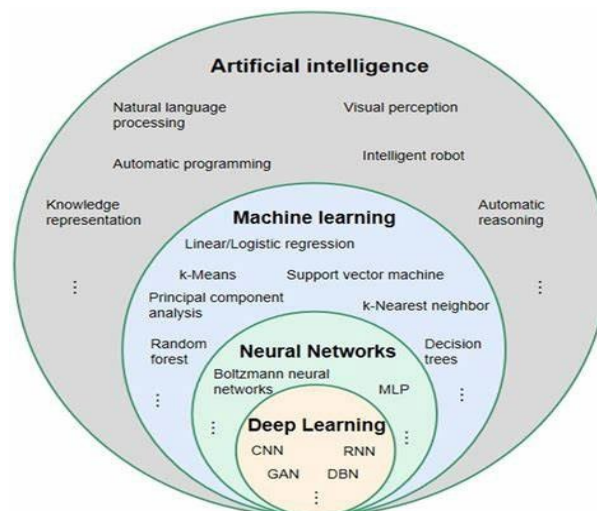
$$AWEI = 4(GREEN - SWIR1) - (0.25 \cdot NIR + 2.75 \cdot SWIR2)$$

2.2.4. Aprendizaje profundo (Deep Learning)

El Deep Learning (DL) es una subcategoría del Aprendizaje Automático (Machine Learning) que utiliza redes neuronales artificiales con múltiples capas de procesamiento para extraer progresivamente características de alto nivel a partir de datos crudos. Como explican LeCun et al. (2015), a diferencia de los métodos estadísticos tradicionales, el DL permite el aprendizaje representacional, donde el modelo identifica automáticamente los patrones relevantes sin necesidad de una extracción manual de características.

En el contexto de la teledetección, el DL ha demostrado una capacidad superior para manejar la complejidad de las imágenes multiespectrales, permitiendo que la red aprenda la relación no lineal entre la reflectancia de las bandas (especialmente la banda NIR).

Figura 6. Deep Learning y Machine Learning dentro de la Inteligencia Artificial.



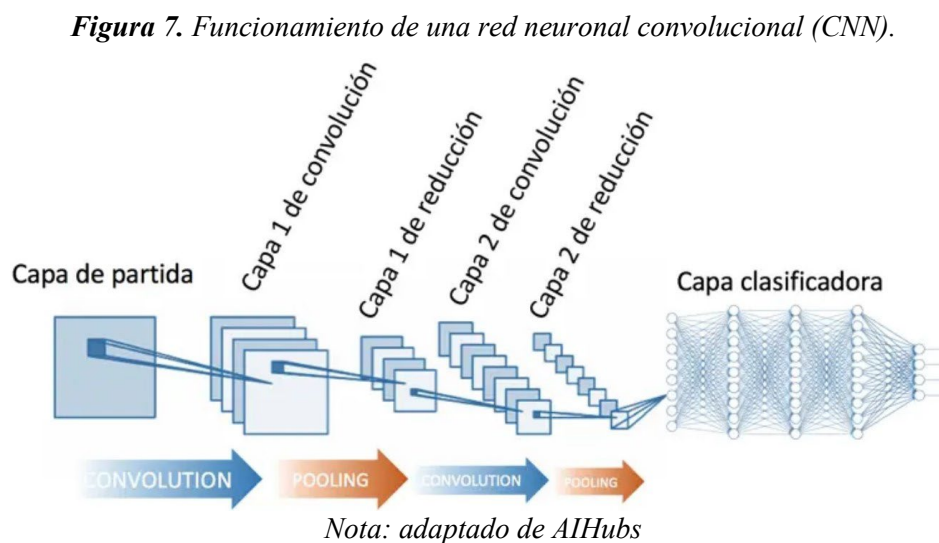
Nota: obtenida de <https://www.researchgate.net>

2.2.4.1. Red neuronal convolucional

Las redes neuronales convolucionales (CNN) representan una categoría avanzada de arquitecturas de Deep Learning diseñadas específicamente para procesar datos con una estructura de rejilla, siendo las imágenes satelitales multiespectrales su aplicación más prominente en la teledetección. Según Lecun et al. (2015), estas redes han impulsado grandes avances revolucionarios en el procesamiento de imágenes al aprovechar las propiedades de las señales naturales mediante cuatro ideas clave: conexiones locales, pesos compartidos, agrupamiento (pooling) y el uso de muchas capas. A diferencia de las redes neuronales densas, las CNN emulan el funcionamiento de la corteza visual humana, analizando patrones locales antes de comprender la estructura global de la imagen.

Las redes neuronales convolucionales tienden a comenzar con un “escáner” de entrada que no está destinado a analizar todos los datos de entrenamiento a la vez. (Lecun et al., 1998).

A continuación, se presenta la **Figura 7** en donde se muestra las capas que conforman una red neuronal convolucional comenzando por la capa de entrada, capas de convolución, capas de reducción, capa clasificadora y finalmente hasta la capa de salida el cual es la encargada de dar la predicción.



a) Capas convolucionales

Es el componente fundamental de la red. Consiste en un conjunto de filtros (kernels) aprendibles que se desplazan a lo largo de la imagen satelital para realizar operaciones matemáticas de producto escalar. Estos filtros permiten detectar características jerárquicas: las capas iniciales identifican bordes y colores, mientras que las capas más profundas reconocen estructuras complejas como la morfología de una cuenca o la delimitación de un reservorio (Goodfellow et al., 2016).

La capa convolucional es el núcleo operativo de la red. En esta etapa, la red aplica un conjunto de filtros o kernels (pequeñas matrices de pesos) que se desplazan sobre la imagen de entrada de 256x256 píxeles para generar "mapas de características".

- **Aprendizaje Jerárquico:** Las capas iniciales de su modelo identifican elementos básicos como bordes de lagunas y texturas de suelo, mientras que las capas profundas logran reconocer estructuras complejas como meandros de ríos o la firma espectral combinada de la vegetación de ribera.
- **Operación Matemática:** Cada neurona en una capa convolucional recibe entradas, las pondera mediante pesos (w_i), suma un sesgo (b) y aplica una función de activación para generar una salida (y).

b) Capas de agrupamiento (Pooling layers)

Las capas de agrupamiento actúan como una etapa de submuestreo que reduce la resolución espacial de los mapas de características generados por las convoluciones. Su implementación es crítica para la eficiencia de un modelo Deep Learning por tres razones fundamentales.

- **Reducción de Dimensionalidad:** Al disminuir el ancho y el alto de los tensores (por ejemplo, de 256x256 a 128x128 mediante un kernel de 2x2), se reduce drásticamente el número de parámetros computacionales.
- **Invariancia a la Traslación:** El agrupamiento, especialmente el Max-Pooling selecciona el valor máximo en cada región de la ventana. Esto permite que la red detecte la presencia de una característica.
- **Aumento del Campo Receptivo:** Al reducir el tamaño de la imagen, las capas convolucionales posteriores pueden "ver" una porción más grande de la imagen original, permitiendo que la red aprenda relaciones contextuales globales.

c) Funciones de activación

La función de activación, también conocida como función de transferencia, establece la relación entre las entradas y salidas de un nodo en una red neuronal. Su introducción de no linealidad es crucial para aplicaciones efectivas de redes neuronales. (Zhang & Morris, 1998).

Las funciones de activación son operadores matemáticos que introducen no linealidad en el modelo, permitiendo que la red aprenda patrones complejos que no podrían resolverse mediante transformaciones lineales simples.

Las funciones más utilizadas en la activación de redes neuronales convolucionales son:

- **Unidad Lineal Rectificada (ReLU):** Es la función estándar en las capas ocultas de su arquitectura de redes convolucionales. Su principal ventaja es que mitiga el problema del desvanecimiento del gradiente (vanishing gradient), permitiendo que redes profundas converjan de manera más rápida y eficiente al mantener un gradiente constante para valores positivos (Nair & Hinton, 2010).

$$f(x) = \max(0, x)$$

- **Leaky Rectified Linear Unit (LeakyReLU):** Introducida inicialmente por Maas et al. (2013), esta variante asigna una pequeña pendiente constante a los valores negativos del umbral. Posteriormente, su efectividad en arquitecturas profundas de visión computacional fue consolidada por He et al. (2015), permitiendo un flujo continuo de gradientes durante la etapa de retropropagación. A diferencia de su predecesora, LeakyReLU no bloquea ni inactiva por completo las señales de entrada negativas, sino que permite el paso de un gradiente pequeño y distinto de cero. Matemáticamente, se define como una función a trozos donde $f(x) = x$ para los valores positivos ($x > 0$), y $f(x) = \alpha \cdot x$ para los valores negativos ($x \leq 0$), siendo α un coeficiente de pendiente negativa constante.
- **Función Sigmoide:** La función sigmoide ha sido un pilar fundamental en el desarrollo del aprendizaje profundo, siendo consolidada por Rumelhart et al. (1986) al introducir el algoritmo de retropropagación debido a su naturaleza continua y derivable. No obstante, investigaciones posteriores como la de Glorot & Bengio (2010) evidenciaron que su uso en capas ocultas de redes profundas genera saturación de gradientes, lo que limita la velocidad de convergencia. Por este motivo, la función sigmoide se reserva principalmente para la capa de salida en problemas de clasificación binaria, donde su capacidad para mapear valores en el rango $[0, 1]$ permite interpretar las activaciones como probabilidades (Goodfellow et al., 2016).

$$f(x) = \frac{1}{1 + e^{-x}}$$

d) Aprendizaje mediante Retropropagación

La Retropropagación es el algoritmo fundamental que permite el entrenamiento de la red mediante la actualización de sus pesos sinápticos. Este proceso se sustenta en el cálculo del gradiente de la función de pérdida con respecto a cada peso del modelo:

- **Paso hacia adelante (Forward Pass):** La imagen satelital (RGB+NIR) atraviesa la red, generando una predicción de segmentación.
- **Cálculo de la Pérdida:** Se cuantifica el error comparando la predicción con la máscara real (Ground Truth).
- **Paso hacia atrás (Backward Pass):** Utilizando la regla de la cadena del cálculo multivariable, el algoritmo propaga el error desde la capa de salida hacia la entrada. El gradiente indica la dirección en la que deben ajustarse los pesos para minimizar el error.
- **Optimización:** Los pesos se actualizan utilizando un optimizador como Adam (Kingma y Ba, 2015), que ajusta la magnitud del cambio basándose en una tasa de aprendizaje (learning rate) predefinida, guiando al modelo hacia el mínimo global de la función de pérdida (Goodfellow et al., 2016).

2.2.4.2. Arquitectura U-Net para segmentación semántica

La arquitectura U-Net, propuesta originalmente por Ronneberger et al. (2015), es una red neuronal convolucional diseñada específicamente para la segmentación semántica de imágenes, donde el objetivo es asignar una categoría a cada píxel individual. Su nombre deriva de su estructura simétrica en forma de "U", la cual permite una transferencia eficiente de información desde las etapas de extracción de características hasta la reconstrucción de la resolución espacial original.

a) Camino de contracción (Encoder/Codificador)

La parte izquierda de la arquitectura constituye el codificador, cuya función es capturar el contexto y extraer características semánticas profundas de la imagen multiespectral.

- **Estructura:** Se compone de bloques repetitivos de dos convoluciones de 3x3 (seguidas de una activación ReLU) y una operación de Max-Pooling de 2x2.
- **Funcionalidad:** En cada nivel, la resolución espacial se reduce a la mitad mientras que el número de canales (filtros) se duplica.

b) Cuello de botella (Bottleneck)

Es la capa más profunda de la red, situada en la base de la "U". En este punto, la imagen ha sido reducida a su mínima expresión espacial, pero posee la máxima profundidad de características. Aquí, la red comprende de manera abstracta basándose en el contexto global de la imagen, independientemente de su ubicación exacta.

c) Camino de Expansión (Decoder / Decodificador)

La parte derecha de la arquitectura se encarga de la localización precisa y la reconstrucción de la máscara binaria.

- **Mecanismo:** Utiliza capas de sobremuestreo (Up-sampling) o convoluciones transpuestas que aumentan gradualmente la resolución espacial mientras reducen la profundidad de los mapas de características.
- **Objetivo:** El decodificador transforma las representaciones abstractas del cuello de botella en un mapa de píxeles que coincide con las dimensiones de la imagen original.

d) Conexiones de Salto (Skip Connections)

Esta es la innovación más trascendental de la U-Net y el factor crítico para su éxito en la teledetección de cuerpos de agua.

2.3. Marco conceptual.

En esta sección se han definidos los términos operativos bajo las cuales se ha desarrollado el trabajo de investigación.

2.3.1. Conceptos de teledetección y análisis espectral

- **Teledetección (sensórica remota):** Ciencia y tecnología que permite obtener información de la superficie terrestre a distancia mediante sensores instalados en plataformas espaciales (satélites) que captan la radiación electromagnética.
- **Imagen multiespectral:** Representación digital compuesta por múltiples bandas que capturan datos en longitudes de onda fuera del espectro visible, como el Infrarrojo Cercano (NIR) y de Onda Corta (SWIR), esenciales para la discriminación hídrica.
- **Pixel:** unidad mínima de información de una imagen satelital que representa un área específica.
- **Banda:** es un rango específico de longitudes de onda electromagnéticas (como luz visible, infrarrojo cercano, etc.) que un sensor capta y registra como una matriz de valores de píxeles.
- **Firma espectral del agua:** comportamiento radiométrico caracterizado por una baja reflectancia en el visible y una absorción casi total en las bandas infrarrojas.
- **Índices espectrales (NDWI, MNDWI, AWEI):** algoritmos aritméticos que combinan bandas específicas para resaltar la presencia de agua y atenuar el ruido de la vegetación, el suelo y las zonas urbanas.
- **Normalización Min-Max:** técnica de preprocesamiento que escala los valores de reflectancia de los píxeles a un rango estándar (generalmente [0 ,1]) para estabilizar el entrenamiento de la red.

2.3.2. Arquitectura y componentes del modelo Deep Learning (U-Net)

- **Kernels (filtros de convolución):** Son matrices numéricas pequeñas (comúnmente de 3×3 píxeles) que se deslizan sobre la imagen de entrada realizando multiplicaciones matemáticas elemento por elemento.
- **Convolución transpuesta (deconvolución):** Es una operación matemática utilizada en redes neuronales para aumentar el tamaño espacial (ancho y alto) de un mapa de características, realizando el proceso inverso a una convolución tradicional.
- **Pooling (agrupación/reducción):** Es una operación de reducción espacial que disminuye el ancho y el alto de las imágenes (generalmente a la mitad) seleccionando solo los valores más importantes de un bloque de píxeles. En U-Net se utiliza específicamente el Max-Pooling de (2×2) , el cual toma únicamente el valor máximo de cada cuadrante.
- **Upsampling (sobremuestreo/escalado hacia arriba):** Es el proceso inverso al pooling, encargado de aumentar las dimensiones espaciales (ancho y alto) de los mapas de características que previamente habían sido reducidos. Se puede realizar mediante interpolación bilineal o convoluciones transpuestas (Transposed Convolutions).
- **Segmentación semántica:** tarea de asignar una etiqueta de clase (agua vs no agua) a cada pixel de la imagen satelital.
- **U-Net:** Arquitectura de red neuronal convolucional simétrica diseñada originalmente para segmentación biomédica, compuesta por un camino de contracción y uno de expansión.

- **Encoder (camino de contracción):** Parte inicial de la red que aplica convoluciones y pooling para extraer características jerárquicas y capturar el contexto global de la imagen.
- **Decoder (camino de expansión):** Parte simétrica que utiliza convoluciones transpuestas para recuperar la resolución espacial original, permitiendo una localización precisa del objeto segmentado.
- **Skip connections (conexiones de salto):** enlaces que concatenan mapas de características del encoder directamente con el decoder, permitiendo que la red recuerde detalles finos de los bordes del agua que suelen perderse durante la compresión de datos.
- **Intersection over Union (IoU):** métrica estadística que cuantifica la precisión del modelo midiendo el solapamiento entre la predicción y el dato real.
- **Bottleneck (Cuello de Botella):** La capa más profunda de la red donde la información está más comprimida y se representan las características más abstractas de la imagen.

2.3.3. Realización, entrenamiento y optimización.

- **Dataset (conjunto de datos):** colección de imágenes satelitales y sus respectivas mascararas utilizadas para entrenar y validar el modelo.
- **Aumentación de datos (data augmentation):** Técnicas de manipulación (rotación, espejado, ajustes de brillo) aplicadas a las imágenes originales para aumentar artificialmente el tamaño del dataset y mejorar la capacidad de generalización del modelo frente a variaciones climáticas o topográficas.
- **Hiperparámetros:** variables externas al modelo que definen el proceso de aprendizaje, tales como la tasa de aprendizaje (Learning rate), el tamaño del lote (batch size) y el número de épocas (epochs).

- **Tasa de aprendizaje:** parámetro que controla la magnitud de los ajustes aplicados a los pesos de la red en cada iteración del optimizador.
- **Optimizador (por ejemplo, Adam):** algoritmo encargado de actualizar los pesos de la red basándose en el cálculo de los gradientes para minimizar la función de pérdida (Kingma y Ba, 2015).
- **Función de pérdida (Loss Function):** Métrica que cuantifica la discrepancia entre la máscara predicha por la red y la máscara real (Ground Truth), como la Entropía Cruzada Binaria (BCE) o el Dice Loss.
- **Retropropagación (Backpropagation):** Algoritmo central que calcula los gradientes de la función de pérdida respecto a cada peso de la red, permitiendo el aprendizaje automático.

2.3.4. Evaluación y métricas de desempeño

- **Intersection over Union (IoU) / Índice de Jaccard:** Métrica que mide el solapamiento entre la predicción y el dato real. Es la métrica estándar para evaluar la precisión en segmentación semántica.
- **Coeficiente de Dice:** Métrica similar al IoU que evalúa la similitud entre dos conjuntos de píxeles, siendo muy robusta ante el desbalance de clases (poca agua respecto al terreno)²⁰.
- **Precisión y Recall:** La precisión mide cuántas de las detecciones del modelo fueron realmente cuerpos de agua, mientras que el *recall* mide cuántos de los cuerpos de agua reales fueron detectados por el modelo²¹.
- **Sobreajuste (Overfitting):** Fenómeno donde el modelo aprende excesivamente los detalles del dataset de entrenamiento, perdiendo la capacidad de realizar detecciones precisas en nuevas imágenes de la provincia de Cangallo.

CAPÍTULO III. MÉTODO DE LA INVESTIGACIÓN.

3.1. Enfoque.

La presente investigación se enmarca dentro de un enfoque estrictamente **cuantitativo**, dado que se fundamenta en la recolección, medición y análisis numérico de datos espaciales para probar las hipótesis planteadas. Este enfoque permite la evaluación objetiva de la capacidad predictiva del modelo de Deep Learning mediante el cálculo de métricas estadísticas estandarizadas, tales como el Índice de Jaccard y la exactitud global, minimizando la subjetividad en la interpretación de la detección de los cuerpos de agua en la topografía andina.

3.2. Alcance.

En cuanto a su nivel, la investigación adopta un alcance **descriptivo-predictivo** de carácter tecnológico y aplicado. Esta investigación se orienta a desarrollar, evaluar y predecir el comportamiento de un modelo computacional complejo en la tarea de segmentación semántica, analizando su rendimiento iterativo para automatizar la identificación de cuerpos de agua frente a escenarios geográficos reales.

3.3. Diseño de investigación.

La presente investigación corresponde a un diseño **experimental**, bajo la categoría de **cuasi-experimento tecnológico** (o experimento de laboratorio computacional).

De acuerdo con las directrices metodológicas de Hernández Sampieri et al. (2014), un estudio adquiere un carácter experimental cuando existe una manipulación intencional y controlada de al menos una variable independiente para analizar sus efectos sobre una o más variables dependientes. En este contexto, la Variable Independiente (el modelo de Deep Learning basado en la arquitectura U-Net y sus configuraciones de hiperparámetros y técnicas de regularización como Data Augmentation) fue activamente manipulada y

calibrada en el laboratorio computacional por el investigador. El objetivo de esta manipulación fue evaluar su impacto directo sobre la Variable Dependiente (la precisión en la segmentación automatizada de cuerpos de agua y la reducción de falsos positivos por ruido orográfico).

El estudio posee una estructura experimental de contraste emparejado, dado que se sometió exactamente al mismo conjunto de prueba aislado (Test Set) a dos estímulos diferenciados: el algoritmo propuesto (U-Net) y el método convencional de teledetección (índice espectral MNDWI), actuando este último como la línea base o grupo de control analítico. La aleatoriedad interna de los procesos de entrenamiento y la posterior validación inferencial mediante las pruebas de Chi-cuadrado de Pearson y McNemar confirman la naturaleza experimental del diseño, demostrando científicamente que los cambios y optimizaciones en las métricas de precisión son una consecuencia causal directa del diseño algorítmico implementado y no de variables intervinientes o del azar.

3.4. Población y muestra.

3.4.1. Población.

Conjunto de imágenes satelitales disponibles en plataformas satelitales (Sentinel-2A) correspondientes a la extensión geográfica de la provincia de Cagallo capturadas durante el año 2024.

3.4.2. Muestra.

Para la presente investigación, se determinó la selección de una muestra **no probabilística por conveniencia**. Este tipo de muestreo se justifica debido a que las unidades de análisis espacial (escenas satelitales y sus respectivos parches) no se eligieron mediante mecanismos puramente aleatorios o probabilísticos, sino bajo criterios de inclusión técnicos, accesibilidad operativa y viabilidad analítica. Específicamente, el

criterio de conveniencia estuvo condicionado por la disponibilidad de imágenes de la misión Sentinel-2A a través de la plataforma Copernicus Browser, seleccionando aquellas capturas correspondientes al año 2024 que cumplieran con el requisito crítico de presentar un porcentaje de cobertura de nubosidad menor al 10% sobre la compleja orografía de la provincia de Cangallo.

A nivel macro, la muestra quedó constituida por 1,327 parches geoespaciales con una dimensión matricial fija de 256×256 píxeles por banda espectral. Esta cantidad de subimágenes y sus dimensiones están técnicamente justificadas por las restricciones de hardware y memoria de acceso aleatorio de video (VRAM) de las unidades de procesamiento gráfico (GPU) utilizadas para el entrenamiento del modelo de Deep Learning. El particionamiento por conveniencia en este formato estandarizado es el requerimiento operativo fundamental para viabilizar el flujo de tensores a través de las capas convolucionales de la arquitectura U-Net. Para el proceso de aprendizaje, este conjunto se subdividió en **1,115** parches (84%) asignados al entrenamiento/validación y **212** parches (16%) aislados estrictamente para la fase de prueba (Test Set).

A nivel micro, para efectos del contraste formal de las hipótesis mediante estadística inferencial (pruebas no paramétricas de Chi-cuadrado de Pearson y de McNemar), la unidad de observación real corresponde al píxel individual evaluado. Dado que la validación científica se ejecuta de forma ciega y con un control estricto sobre el Test Set (los 212 parches independientes remanentes), el tamaño neto de la muestra estadística (N) se deriva del cálculo geométrico de las dimensiones matriciales:

$$N = 212 \text{ parches} \times 256 \text{ píxeles} \times 256 \text{ píxeles} = 13,893,632 \text{ píxeles}$$

Por consiguiente, aunque el origen de las imágenes responde a un criterio no probabilístico por conveniencia, la validación final del modelo se sustenta en un universo

masivo de 13,893,632 observaciones nominales binarias. Este volumen masivo a escala de píxeles dota a las pruebas estadísticas de una potencia analítica máxima, asegurando que las inferencias y el rechazo de las hipótesis nulas.

3.5. Hipótesis.

3.5.1. Hipótesis general.

La implementación de un modelo de Deep Learning basado en la arquitectura U-Net influye positivamente en la detección automatizada de cuerpos de agua en la provincia de Cangallo, logrando un rendimiento de grado cartográfico.

3.5.2. Hipótesis específicas.

- a) El modelo U-Net alcanza una precisión superior al 85% en la detección de cuerpos de agua.
- b) El modelo U-Net alcanza un índice IoU superior a 0.85, evidenciando alta capacidad en la segmentación.
- c) Existe una diferencia estadísticamente significativa ($p < 0.05$) a favor del modelo U-Net frente al índice MNDWI en la detección de cuerpos de agua.

3.6. Operacionalización de variables.

3.6.1. Definición conceptual y operacional.

En la **Tabla 3.** se realiza la definición conceptual y operaciones de las variables identificadas en este estudio.

Tabla 3. Definiciones de las variables

Variable	Definición conceptual	Definición operacional
V. Independiente: Modelo de Deep Learning	Sistema computacional basado en redes neuronales artificiales compuestas por múltiples capas de procesamiento (de ahí el término "profundo")	En la práctica, esta variable se manipula y aplica mediante la implementación de un script computacional en el lenguaje Python (haciendo uso de bibliotecas como Keras y TensorFlow)
V. Dependiente: Detección de cuerpos de agua	Capacidad del sistema para asignar una etiqueta binaria (agua/no agua) a cada píxel, identificando con precisión la forma y extensión de las masas hídricas.	Cálculo matemático de la coincidencia entre la máscara generada por el modelo y la máscara de referencia (Ground Truth) mediante algoritmos de validación métrica.

Nota: Elaboración propia.

3.6.2. Operacionalización de variable independiente y dependiente

En la **Tabla 4.** y **Tabla 5.** se ha operacionalizado la variable independiente (modelo Deep Learning) y la variable dependiente (Detección de cuerpos de agua), respectivamente.

Tabla 4. Operacionalización de la Variable Independiente.

Variable	Dimensiones	Indicadores	Técnica/ Instrumento	Escala	Unidad de media
V. I: Modelo de Deep Learning	Arquitectura del modelo	Tipo de red (U-Net)	Revisión del modelo	Nominal	Tipo
	Entrenamiento	Número de épocas (epochs)	Registro de entrenamiento	Razón	Número
		Función de pérdida (Loss)	Software (TensorFlow)	Razón	Valor
	Configuración	Número de capas	Análisis del modelo	Razón	Número
		Tasa de aprendizaje (Learning rate)	Configuración del modelo	Razón	Valor
	Procesamiento de datos	Aplicación de Data Augmentation	Registro del Proceso	Nominal	Si / No

Nota: Elaboración propia.

Tabla 5. Operacionalización de la Variable Dependiente.

Variable	Dimensiones	Indicadores	Técnica/ Instrumento	Escala	Unidad de media
V. D: Detección de cuerpos de agua	Precisión global	Accuracy (exactitud)	Matriz de confusión	Razón	%
		IoU (Intersection over Union)	Evaluación del modelo	Razón	%
	Desempeño de segmentación	Precision (precisión positiva)	Matriz de confusión	Razón	%
		Recall (sensibilidad)	Matriz de confusión	Razón	%
	Identificación	Número de cuerpos detectados correctamente	Análisis espacial (QGIS)	Razón	Cantidad
	Comparación	Diferencia con índice MNDWI	Prueba de McNemar	Nominal / Inferencial	

Nota: Elaboración propia.

3.7. Técnicas e instrumentos.

3.7.1. Técnicas.

3.7.1.1. Técnicas de recolección y procesamiento.

Las técnicas representan el conjunto de procedimientos metódicos utilizados para obtener y preparar la información necesaria para el entrenamiento del modelo U-Net.

- **Teledetección pasiva (Remote sensing):** Técnica fundamental para la adquisición de datos multiespectrales sin contacto físico. Se basa en la captura de la reflectancia de la superficie terrestre mediante los sensores MSI de Sentinel-2.
- **Procesamiento digital de imágenes:** conjunto de algoritmos aplicados para la corrección y preparación de los datos incluye normalización radiométrica (técnica para la estandarización de valores de los píxeles en un rango de 0 a 1) y fragmentación espacial (Tiling, es una técnica de subdivisión de escenas de gran formato en parches de 256x256 píxeles para su procesamiento).

- **Aprendizaje supervisado:** Técnica de entrenamiento de la red neuronal mediante el uso de pares de datos: la imagen satelital original y su respectiva máscara de referencia (Ground Truth), permitiendo que el modelo aprenda a identificar patrones hídricos de forma iterativa.
- **Aumentación de datos (Data Augmentation):** Técnica para incrementar artificialmente la variabilidad de la muestra mediante transformaciones geométricas (rotaciones, espejados) y fotométricas, mejorando la capacidad de generalización del modelo.

3.7.2. Instrumentos.

Los instrumentos son las herramientas lógicas y físicas que hacen posible la ejecución de las técnicas mencionadas.

3.7.2.1. Fuentes de información (sensores remotos).

Se utilizarán como instrumentos de captura de datos las misiones satelitales Sentinel-2A (sensor MSI). El sensor proporciona reflectancia en el fondo de la atmósfera (BOA), ideal para evitar distorsiones por neblina en los Andes.

3.7.2.2. Instrumentos de software (entorno lógico).

El procesamiento se realizará íntegramente en un entorno de desarrollo basado en Python 3.10, utilizando las siguientes librerías especializadas:

- **Google Earth Engine (GEE) API:** Instrumento para la descarga y filtrado masivo de imágenes satelitales por fecha y nubosidad.
- **TensorFlow / Keras:** Framework principal para la construcción, entrenamiento y despliegue de la arquitectura U-Net.

- **Rasterio y GDAL:** Librerías para la manipulación de archivos geospaciales en formato GeoTIFF.
- **Arcgis Pro:** Sistema de Información Geográfica (SIG) utilizado para la inspección visual, etiquetado manual de máscaras y validación final de mapas vectoriales.

3.7.2.3. Instrumentos de hardware (entorno físico).

Dada la alta demanda computacional de las Redes Neuronales Convolucionales, se empleará:

- **Unidad de Procesamiento Gráfico (GPU):** NVIDIA T4 o RTX 3060 (mínimo 8GB VRAM) para la aceleración del cálculo de tensores y retropropagación.
- **Memoria RAM:** Mínimo 16 GB para la gestión de arreglos multidimensionales en memoria.
- **Almacenamiento:** Discos de estado sólido (SSD) para una lectura rápida de los parches de imagen durante el entrenamiento.

3.7.2.4. Validez y confiabilidad de los instrumentos.

- **Validez de los Datos:** Se garantiza mediante el uso de productos satelitales calibrados (Nivel 2A) con corrección atmosférica. La validez de las máscaras de referencia se confirma mediante verificación visual en imágenes de alta resolución (Google Earth/Planet) y el cálculo de índices espectrales (MNDWI) como línea base.
- **Confiabilidad del Modelo:** Se establece mediante la técnica de Validación Cruzada (Cross-Validation) y el monitoreo de las curvas de pérdida (Loss) y precisión (Accuracy) para detectar y corregir el sobreajuste (overfitting).

3.8. Técnicas estadísticas para el procesamiento de la información.

Esta sección describe las herramientas matemáticas y estadísticas que se emplearán para analizar los datos obtenidos en la fase experimental computacional, permitiendo la validación técnica de las hipótesis planteadas.

3.8.1. Métricas de la evaluación del desempeño (segmentación).

Para cuantificar la precisión de la arquitectura U-Net en la provincia de Cangallo, se utilizarán métricas basadas en la Matriz de Confusión a nivel de píxel. Cada píxel clasificado se categoriza como: Verdadero Positivo (VP), Verdadero Negativo (VN), Falso Positivo (FP) o Falso Negativo (FN).

- **Precisión (Precision):** mide la exactitud de las predicciones positivas (cuantos píxeles del agua detectada realmente es agua)

$$Precision = \frac{VP}{VP + FP}$$

- **Exhaustividad (Recall):** Mide la capacidad del modelo para encontrar todos los píxeles de agua existentes.

$$Recall = \frac{VP}{VP + FN}$$

- **F1-Score:** Es la media armónica entre precisión y exhaustividad, proporcionando un balance entre ambas.

$$Recall = 2 * \frac{Precision * Recall}{Precision + Recall}$$

3.8.2. Índice de Jaccard (IoU).

El indicador principal de éxito será el IoU. Esta métrica es el estándar en segmentación semántica, ya que penaliza los falsos positivos y evalúa el solapamiento exacto entre la máscara predicha por la red (A) y la máscara de referencia (B).

$$IoU(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{VP}{VP + FP + FN}$$

Un valor que sea superior al 0.85 confirmará la validez del modelo Deep Learning utilizada para las detecciones de cuerpos de agua en la zona de estudio.

3.8.3. Coeficiente de coincidencia de Dice

Complementariamente, se utilizará el coeficiente de Dice, el cual es más sensible al desbalance de clases (cuando los cuerpos de agua ocupan un área muy pequeña respecto al terreno total de la imagen satelital).

$$Dice = \frac{2 \cdot VP}{2 \cdot VP + FP + FN}$$

3.8.4. Estadística inferencial y prueba de hipótesis

La estadística inferencial en este estudio constituye el mecanismo matemático que permite extrapolar los resultados obtenidos en el entorno computacional hacia conclusiones científicamente válidas. Su propósito es demostrar, con un nivel de confianza del 95% ($\alpha = 0.05$), que el rendimiento del modelo Deep Learning no es producto del azar, sino de un aprendizaje algorítmico efectivo. Para contrastar las hipótesis planteadas sobre el conjunto ciego de prueba (Test Set), se aplicarán dos pruebas estadísticas fundamentales:

3.8.4.1. Prueba de Chi-Cuadrado de Pearson (validación de asociación)

Para la validación de la hipótesis general de la investigación, se seleccionó la prueba estadística no paramétrica de **Chi-cuadrado de Pearson de Independencia**. En el marco

metodológico del modelamiento predictivo mediante Deep Learning aplicado a la teledetección, esta prueba inferencial constituye el estándar científico para evaluar la existencia de una relación o asociación significativa entre dos variables cualitativas de tipo nominal dicotómico.

Específicamente, el análisis se ejecuta cruzando la variable de Predicción Algorítmica del Modelo U-Net (cuyas categorías discretas son "Agua" y "No Agua") frente a la variable de Verdad de Terreno o Ground Truth (cuyas categorías homólogas son "Agua" y "No Agua"), estructurando los datos en una tabla de contingencia clásica de 2×2 (matriz de confusión global).

Las hipótesis que formula esta prueba son:

Hipótesis nula (H0): Las clasificaciones geoespaciales generadas de forma automatizada por la red neuronal convolucional U-Net y las etiquetas reales del terreno son variables estadísticamente independientes. Bajo este supuesto, las inferencias del modelo no poseen un sustento de aprendizaje físico-espectral, siendo el rendimiento del algoritmo un mero producto del azar o la aleatoriedad.

Hipótesis alterna (H1): Existe una dependencia y asociación matemática altamente significativa entre las predicciones del modelo U-Net y la presencia real de los cuerpos de agua en la superficie. La aceptación de esta hipótesis demuestra formalmente que la Inteligencia Artificial asimiló patrones jerárquicos contextuales genuinos para la discriminación hídrica.

El cálculo del estadístico inferencial se basa en cuantificar la magnitud de la discrepancia cuadrática entre las frecuencias observadas (O_{ij}) recopiladas directamente en la matriz de confusión del conjunto de prueba y las frecuencias teóricas esperadas (E_{ij}) que deberían registrarse si la hipótesis nula de independencia absoluta fuera estrictamente verdadera. La ecuación general que rige el estadístico se expresa de la siguiente manera:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Donde r representa el número de filas y c el número de columnas de la matriz (2×2), lo que determina de forma invariable un grado de libertad $gl = (r - 1)(c - 1) = 1$. Las frecuencias esperadas para cada celda de la tabla de contingencia se derivan matemáticamente a partir del producto de sus respectivos marginales de fila (F_i) y columna (C_j), divididos entre el tamaño total de la muestra estadística (N):

$$E_{ij} = \frac{F_i \cdot C_j}{N}$$

La validez matemática de la prueba de Chi-cuadrado requiere el cumplimiento del supuesto de frecuencias esperadas robustas, estipulando que ninguna celda teórica debe registrar un valor inferior a 5 observaciones. Debido a que la unidad de observación en esta fase de la tesis corresponde al píxel individual evaluado sobre el conjunto de parches de prueba ciego (Test Set), el tamaño de la muestra estadística asciende a un volumen masivo de $N = 13,893,632$ píxeles, satisfaciendo con creces cualquier restricción distributiva.

No obstante, Lin et al. (2011) advierte que ante muestras de magnitud macroscópica ($N \rightarrow \infty$), el estadístico χ^2 sufre de una hipersensibilidad asintótica conocida como la paradoja del tamaño de la muestra. Bajo este escenario de macrodatos, cualquier desviación física minúscula o irrelevante entre lo observado y lo esperado infla exponencialmente el valor calculado de Chi-cuadrado, provocando que el valor de probabilidad asociado (p -value) converja virtualmente a cero de forma matemática artificial.

A fin de mitigar el sesgo inflacionario intrínseco del volumen masivo de píxeles y evaluar con rigor científico la fuerza de asociación real y neta entre el modelo propuesto y el

terreno, se determinó metodológicamente normalizar el estadístico mediante el cálculo del coeficiente de V de Cramer (V) (el cual equivale matemáticamente al coeficiente Phi ϕ para tablas de 2×2). Su formulación se define como:

$$V = \sqrt{\frac{\chi^2}{N \cdot \min(r - 1, c - 1)}} = \sqrt{\frac{\chi^2}{N}}$$

Este coeficiente actúa como un indicador acotado estrictamente en un rango de 0 a 1, donde 0 representa una independencia absoluta y 1 una correspondencia o asociación perfecta. El análisis de los resultados no se limitará, por tanto, a la mera significancia del p -value, sino a la interpretación del tamaño del efecto bajo las escalas estandarizadas de Cohen. Un coeficiente $V > 0.50$ clasificará la relación como una asociación muy alta, sirviendo de sustento inferencial definitivo para convalidar que la precisión global y el índice de superposición geográfica (IoU) del modelo U-Net responden a una legítima y sobresaliente capacidad de segmentación geomorfológica.

3.8.4.2. Prueba no paramétrica de McNemar (comparación de modelos)

Se utilizará para contrastar la Hipótesis Específica 3, la cual plantea una diferencia estadísticamente significativa a favor de la arquitectura U-Net frente al índice espectral tradicional MNDWI. Dado que ambos métodos de clasificación evaluarán exactamente los mismos píxeles espaciales (muestras emparejadas o dependientes), la prueba de McNemar es el estándar estadístico adecuado, ya que se centra en medir las discrepancias en las tasas de error.

Para su ejecución matemática, los resultados de ambos clasificadores se aplanarán a vectores unidimensionales y se construirán dentro de una tabla de contingencia de 2×2 .

Esta matriz no evalúa los aciertos simples, sino que aísla cuatro escenarios operativos:

- **n00:** Casos donde ambos métodos (U-Net y MNDWI) aciertan en la detección

- **n11:** Casos donde ambos métodos fallan.
- **n01 (discrepancia a favor de U-Net):** Píxeles donde el modelo U-Net clasifica correctamente, pero el MNDWI falla.
- **n10 (Discrepancia a favor de MNDWI):** Píxeles donde el índice MNDWI clasifica correctamente, pero la red U-Net falla.

El estadístico de McNemar se calculará sobre las discrepancias (n01 y n10) mediante una aproximación Chi-Cuadrado. Si el valor p resultante es inferior a 0.05, se rechazará la hipótesis nula de que ambos modelos tienen un desempeño equivalente, permitiendo demostrar estadísticamente la superioridad tecnológica de la inteligencia artificial frente a los falsos positivos (como sombras orográficas) que el método tradicional no logra mitigar.

3.8.5. Herramientas de procesamiento estadístico

El procesamiento de estas métricas se realizará de forma automatizada mediante las librerías a las que se le harán un llamado en el script generado para la predicción del modelo:

- **Librería Scikit-learn (Python):** Para la generación de matrices de confusión y reporte de métricas de clasificación.
- **NumPy y Pandas:** Para la gestión de grandes volúmenes de datos matriciales provenientes de las imágenes multiespectrales y sus mascarar binarias.
- **Matplotlib / Seaborn:** Para la visualización de resultados mediante curvas de precisión-exhaustividad (P-R) y curvas ROC.

3.9. Desarrollo del trabajo de tesis.

Para la ejecución de la presente investigación, se implementará un modelo de Aprendizaje Profundo (Deep Learning) fundamentado en la arquitectura U-Net. Esta red se distingue por su estructura simétrica de codificación y decodificación, diseñada específicamente para la segmentación semántica de alta precisión. La arquitectura integra una ruta de contracción (encoder) para la extracción de características abstractas y una ruta de expansión (decoder) para la recuperación de la resolución espacial, vinculadas mediante conexiones de salto (skip connections). Estas conexiones son críticas para preservar y restaurar los detalles topológicos y texturales de las imágenes multiespectrales, mitigando la pérdida de información durante el proceso de submuestreo.

El desarrollo metodológico se ha estructurado en cuatro fases fundamentales, en las cuales se detallan de forma exhaustiva los procesos de preprocesamiento, diseño, optimización y validación. El objetivo final es el entrenamiento del modelo bajo métricas de rendimiento rigurosas (como el Índice de Jaccard y el F1-Score), garantizando una alta capacidad de generalización al realizar inferencias y predicciones sobre imágenes satelitales inéditas, asegurando así la robustez del sistema en condiciones geográficas diversas.

3.9.1. Fase I: Adquisición y preprocesamiento de datos geoespaciales.

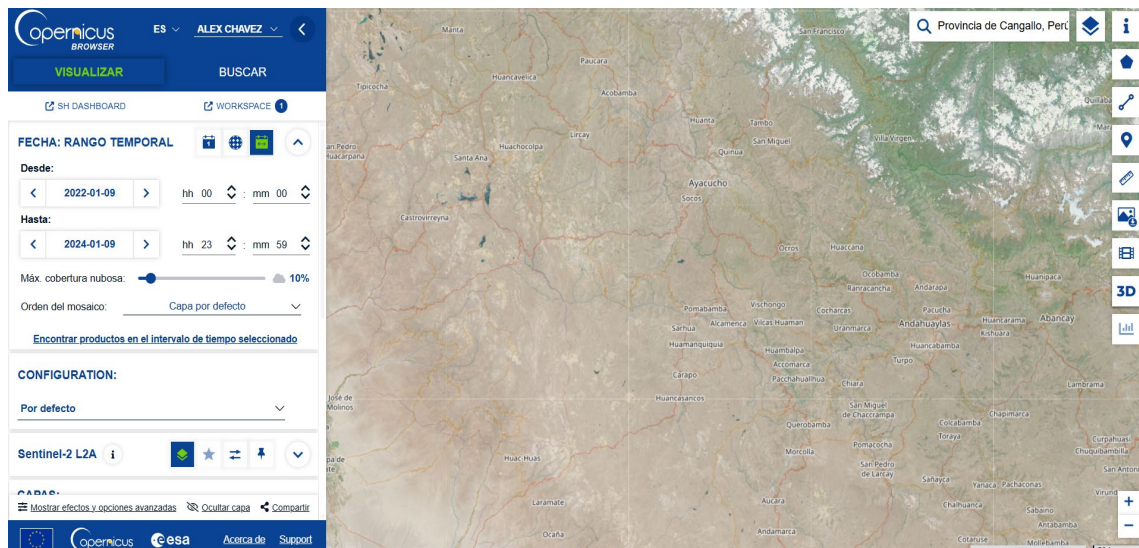
Esta fase constituye el pilar fundamental de la investigación, donde se consolida el dataset de las imágenes multiespectrales y sus mascarar. El flujo de trabajo se ha optimizado para la misión Sentinel-2A (sensor MSI) y un procesamiento de datos selectivo.

3.9.1.1. Descarga de datos geoespaciales.

La obtención de datos se realiza a través Copernicus Browser del ecosistema Copernicus Data Space Ecosystem (CDSE), extrayendo escenas de la misión Sentinel-2A:

- **Selección de escena y periodo temporal:** como escena se ha localizado la provincia de Cangallo y se han descargado gránulos correspondientes al año 2024, lo que garantiza una representatividad multitemporal de los cuerpos de agua.
- **Criterios de calidad:** Se aplicó un filtro estricto de corrección de nubosidad al 10%, asegurando que la señal espectral de los cuerpos de agua no se vea interferida por aerosoles o nubes densas.
- **Descarga:** Se obtienen los archivos comprimidos que contienen las bandas individuales en formato GeoTIFF de 16 bits.

Figura 9. Descarga de granula de Sentinel-2A des de Copernicus Browser.



Nota: elaboración propia

3.9.1.2. Selección y composición de bandas en ArcGis Pro.

La elección de las bandas B3 (Verde), B4 (Rojo), B8 (Infrarrojo Cercano - NIR) y B11 (Infrarrojo de Onda Corta - SWIR) de Sentinel-2A, junto con la exclusión deliberada de la banda azul, responde a criterios estrictos de física de la radiación, firmas espectrales e ingeniería de características para Inteligencia Artificial.

La exclusión deliberada de la banda azul (B2) en el diseño de características espectrales para el modelo U-Net se justifica por razones estrictamente físicas de la radiación y

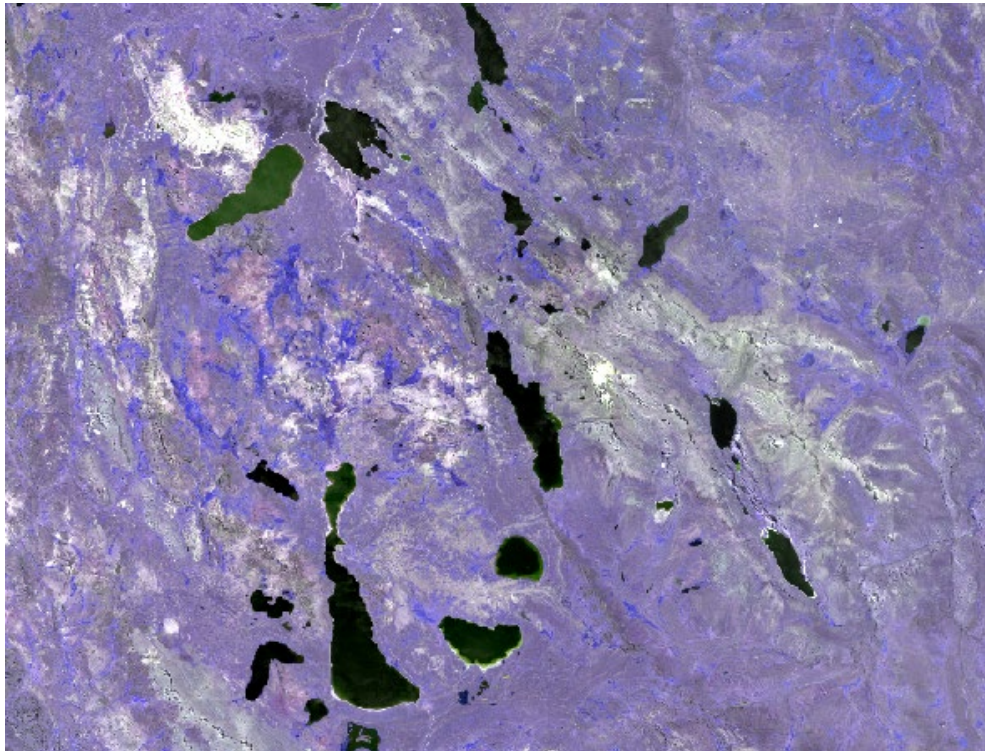
control de ruido en teledetección. Al operar en la longitud de onda más corta del espectro visible ($\approx 490\text{ nm}$), la banda azul es altamente vulnerable a la **dispersión atmosférica de Rayleigh**, un fenómeno óptico donde los gases y aerosoles distorsionan la señal y saturan los píxeles con ruido radiométrico y neblina de fondo. Esta inestabilidad se agrava drásticamente sobre la compleja topografía de la provincia de Cagallo, donde las laderas andinas proyectan profundas sombras orográficas; debido a la luz difusa del cielo (skylight), las sombras no son completamente oscuras en la banda azul, introduciendo firmas espectrales homólogas que confunden a los clasificadores. Por lo tanto, prescindir de este canal impidió que la red neuronal asimilara patrones espaciales e iluminaciones defectuosas, optimizando la capacidad convolucional de las capas para concentrarse únicamente en la absorción hídrica pura y discriminando con precisión matemática una laguna real de una sombra proyectada.

Una vez que se han descargado los datos, se deben importar al software ArcGis Pro para su estructuración y procesamiento.

Se utiliza la herramienta “*Composite Bands*” para unificar los canales necesarios para la detección de agua. En esta investigación se ha dado por conveniente seleccionar las bandas:

- **B3 (verde):** Tiene una alta reflectancia en cuerpos de agua.
- **B4 (Red):** Fundamental para la delimitación morfológica.
- **B8 (NIR):** Banda crítica por su alta absorción en agua.
- **B11 (SWIR):** Esencial para discriminar humedad en el suelo y sombras topográficas.

Figura 10. Imagen multiespectral obtenida por la composición de bandas en ArcGIS Pro



Nota: Elaboración propia

3.9.1.3. Generación de máscaras (Ground Truth).

La construcción de las máscaras binarias de referencia (Ground Truth) constituyó una etapa metodológica crítica para el éxito del aprendizaje supervisado de la red U-Net. Para garantizar que las etiquetas estuvieran completamente libres de contaminación por nubes y falsos positivos causados por sombras orográficas, se ejecutó un flujo de trabajo especializado dentro del entorno ArcGIS Pro (versión 3.7) estructurado bajo tres procedimientos técnicos secuenciales:

1) Selección y recorte geográfico de gránulos por nubosidad

El insumo inicial consistió en las escenas satelitales Sentinel-2A de la provincia de Cangallo. Debido a que la presencia estacional de cobertura nubosa impide la correcta discriminación espectral y visual de las lagunas o "ccochas" altoandinas, se realizó un proceso de filtrado y control de calidad visual sobre los gránulos de las imágenes.

Se identificaron y recortaron vectorialmente aquellas porciones o subregiones del gránulo que presentaban condiciones óptimas de transmitancia atmosférica (libres de nubes o con nubosidad menor al 10%). Este procedimiento de segmentación geográfica previa garantizó que el espacio de búsqueda visual estuviera completamente despejado, viabilizando una delimitación geométrica precisa y evitando que la red neuronal convolucional asimilara patrones de ruido blanco (nubes) como parte de la cobertura del suelo.

2) Digitalización vectorial mediante el panel “Label Objects for Deep Learning”

Con las porciones de las imágenes satelitales depuradas y cargadas en el mapa, se accedió a la pestaña Imagery y se activó el panel especializado Label Objects for Deep Learning (perteneciente a la extensión Spatial Analyst / Image Analyst de ArcGIS Pro). Esta herramienta proporciona un entorno optimizado para el esquema de clasificación supervisada en segmentación semántica pixel a pixel.

A través de este panel, se configuró un esquema de asignación de clase binaria dicotómica estructurada de la siguiente manera:

- **Clase agua (código de valor: 1):** Polígonos digitalizados manualmente de forma exhaustiva sobre los espejos de agua de las lagunas y reservorios. La fotointerpretación directa mediante la visualización de imágenes satelitales de alta resolución y las imágenes satelitales multiespectrales permitió asegurar que los perímetros trazados correspondieran netamente a recursos hídricos, excluyendo de forma absoluta las sombras de laderas andinas que suelen poseer firmas espectrales homólogas.

- **Clase terreno/Fondo (código de valor: 0):** Todo el entorno circundante que compone la escena (suelo desnudo, bofedales, vegetación, afloramientos rocosos y sombras orográficas).

La digitalización manual utilizando este módulo nativo de ArcGIS Pro permitió estructurar un Feature Class de polígonos donde cada elemento vectorial quedó indexado automáticamente con su respectivo identificador de clase binaria (ClassValue).

3) Rasterización, sincronización de grilla y binarización de la máscara

Una vez concluido el trazado vectorial de la totalidad de cuerpos de agua en las zonas de estudio, las geometrías de los polígonos se sometieron al proceso de rasterización para transformarlas en matrices numéricas compatibles con los tensores de Python.

Para este fin, se ejecutó la herramienta de conversión Polygon to Raster, configurando las propiedades del entorno de geoprocésamiento (Environments) para garantizar una sincronización espacial perfecta:

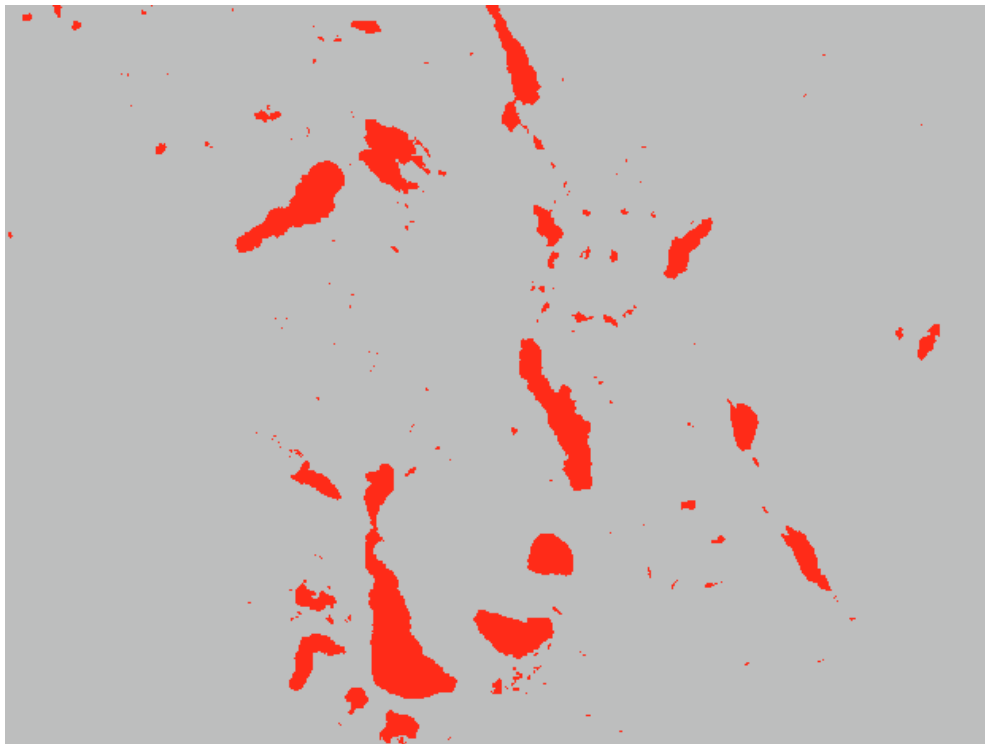
- Se seleccionó el campo de la clase binaria como el valor del píxel de salida.
- Se definió el tamaño de celda (Cell Size) rígidamente en 10 metros, preservando la resolución nativa del sensor MSI de Sentinel-2A.
- Se activó la opción de Ráster de Ajuste (Snap Raster) apuntando a la imagen satelital original. Esto forzó a que la grilla de la máscara generada se alineara perfectamente, píxel por píxel y esquina con esquina, con las matrices multiespectrales de entrada del modelo.

Finalmente, mediante el uso de operadores condicionales en el álgebra de mapas, la matriz provincial continua fue binarizada de forma estricta (valores 1 para el agua y 0

para el fondo), quedando lista para el posterior particionamiento ortogonal en parches homogéneos de 256x256 píxeles requeridos por la arquitectura U-Net.

En la **Figura 11** se muestra la máscara binaria de un recorte de granulo multiespectral, en esta figura los polígonos de color rojo (valor 1) muestran las lagunas delimitadas y el fondo que se muestra gris es el terreno circundante (valor 0).

Figura 11. Mascara binaria de la imagen satelital multiespectral obtenida en ArcGis Pro



Nota: Elaboración propia

3.9.1.4. Exportación de datos de entrenamiento (Tiling).

Este es el puente entre el SIG y el Deep Learning. Se utiliza la herramienta Export Training Data for Deep Learning del software ArcGis Pro:

- **Tile Size (Patch size):** Se configura a 256 x 256 píxeles.
- **Stride (Zancada):** Se establece un desplazamiento de 128 píxeles para generar solapamiento (overlap), lo cual aumenta el tamaño de la muestra.

- **Meta Data Format:** la imagen multiespectral y las máscaras se exportan en carpetas separadas en formato .tif.

3.9.1.5. Filtrado de datos y normalización diferida

Después de proceso de fragmentación, se realizó una auditoria de la calidad de la dataset tanto del entrenamiento y validación, así como el de prueba.

- **Limpieza de máscaras:** Se detectó una alta incidencia de parches con fondo negro (sin presencia de agua). Estos fueron eliminados, conservando únicamente el 10% del total de muestras para el entrenamiento, optimizando así el balance de clases del modelo.
- **Normalización en Python:** Durante el preprocesamiento en ArcGIS no se normalizaron los datos. La normalización se ejecutó directamente en el entorno de Python durante la fase de lectura de datos, transformando las imágenes en matrices mediante la librería NumPy para estabilizar los gradientes durante el entrenamiento de la red U-Net.

3.9.2. Fase II: Diseño y configuración de la arquitectura U-Net.

Esta fase define la estructura lógica y matemática del "cerebro" del sistema. Se ha seleccionado la arquitectura U-Net, un modelo de red neuronal convolucional (CNN) totalmente simétrico, diseñado para realizar una segmentación semántica precisa, asignando una etiqueta de clase a cada píxel individual de la imagen satelital.

La implementación del modelo se realizó en el entorno de investigación Google Colab, aprovechando su infraestructura de computación en la nube de alto rendimiento.

- **Aceleración GPU:** Se emplearon Unidades de Procesamiento Gráfico (como NVIDIA T4) para paralelizar las operaciones matriciales de los tensores, reduciendo los tiempos de entrenamiento.
- **Librerías especializadas:** El código fue desarrollado en Python utilizando el framework TensorFlow/Keras. Se emplearon librerías auxiliares como NumPy para la manipulación de arreglos multidimensionales y Matplotlib para el monitoreo de las curvas de aprendizaje.

3.9.2.1. Capa de entrada (Input Layer)

De acuerdo con el procesamiento de las imágenes satelitales multiespectrales y sus máscaras, el modelo recibe un tensor de entrada con dimensiones de (256, 256, 4).

- **Canales Espectrales:** El modelo integra las bandas B3 (Verde), B4 (Rojo), B8 (NIR) y B11 (SWIR) de Sentinel-2A. La inclusión de la banda NIR y SWIR es fundamental para que la red aprenda la firma de absorción característica del agua.
- **Resolución espacial:** 256 x 256 píxeles, dimensión estándar para optimizar el uso de la memoria VRAM de la GPU sin perder contexto geográfico.

Figura 12. Implementación de la capa de entrada del modelo U-Net

```
import tensorflow as tf
from tensorflow.keras import layers, models, regularizers
from tensorflow.keras import backend as K
import numpy as np

ANCHO_IMG = 256 #esto se ha colocado por que ya no se ha corrido la lectura de imagenes
ALTO_IMG = 256 #esto se ha colocado por que yano se ha corrido la lectura de imagenes

def construccion_unet(n_classes=1, img_height=256, img_width=256, img_channels=3, use_batch_norm=True, dropout_rate=0.3, a
    inputs = layers.Input((img_height, img_width, img_channels))

# Función auxiliar para bloques de convolución
def conv_block(input_tensor, num_filters):
    # Primera Convolución
    x = layers.Conv2D(num_filters, (3, 3), padding="same", kernel_regularizer=kernel_regularizer)(input_tensor)
    if use_batch_norm:
        x = layers.BatchNormalization()(x) # Estabiliza los pesos
    x = activation_layer(x) # Usar la capa de activación dinámica

# Dropout (Opcional: ponerlo aquí o después de la segunda conv)
if dropout_rate > 0:
    x = layers.Dropout(dropout_rate)(x) # Combate el Overfitting
```

Nota: elaboración propia

3.9.2.2. Estructura del camino de contracción (Encoder).

características, reduciendo progresivamente la dimensionalidad espacial de la imagen de entrada (256x256 píxeles) mientras aumenta la profundidad de los mapas de características.

- **Operación de convolución:** Se implementaron bloques de contracción. Cada bloque consta de dos capas consecutivas de convolución (Conv2D) con un tamaño de kernel de 3x3 y función de activación LeakyReLU con una pendiente negativa de 0.1, la cual introduce no linealidad al modelo permitiendo el aprendizaje de patrones complejos.
- **Regularización:** Se aplicó la técnica de padding tipo 'same' para mantener las dimensiones espaciales tras cada convolución.
- **Reducción de dimensiones:** Al finalizar cada bloque, se empleó una capa de agrupamiento máximo (MaxPooling2D) con una ventana de 2x2, reduciendo la resolución de la imagen a la mitad para abstraer la información geométrica.
- **Filtros:** El número de filtros se duplica en cada profundidad (32, 64, 128, 256), permitiendo capturar desde bordes nítidos de lagunas hasta patrones texturales complejos de sombras topográficas.

3.9.2.3. Cuello de botella (Bottleneck).

Es la capa más profunda de la red donde la información geoespacial está más comprimida. Se ha definido una trayectoria de contracción compuesta por bloques convolucionales jerárquicos con las siguientes operaciones secuenciales:

- Aquí se utilizan 512 filtros de convolución.

- Para mitigar el riesgo de sobreajuste (overfitting), se incorporó una capa de Dropout con una tasa de 0.28, desactivando aleatoriamente un porcentaje de las neuronas durante el entrenamiento.

3.9.2.4. Estructura de camino de expansión (Decoder).

Su función es recuperar la resolución espacial original para generar la máscara de segmentación precisa.

- **Sobremuestreo (Upsampling):** Se utilizaron capas de deconvolución o transposición (Conv2DTranspose) con stride de 2x2, que duplican las dimensiones de los mapas de características.
- **Conexiones residuales (skip connections):** Se implementaron conexiones de salto que concatenan (Concatenate) los mapas de características del encoder con los del decoder en el mismo nivel de resolución. Esta técnica es crítica, ya que permite al modelo recuperar la información de ubicación espacial perdida durante la contracción, mejorando la precisión en la delimitación de los bordes de los cuerpos de agua.

3.9.2.5. Capa de salida y función de activación.

La capa final reduce los 32 filtros del último nivel a un solo canal de salida.

- **Activación:** Se aplica la función sigmoide.
- **Interpretación:** El resultado es un mapa de probabilidad donde cada píxel tiene un valor entre 0 (tierra) y 1 (agua). Un umbral de 0.5 servirá para generar la máscara binaria final.

La arquitectura de la red neuronal convolucional U-Net implementada para la segmentación de cuerpos hídricos presenta un diseño simétrico codificador-

decodificador, cuyo flujo de procesamiento algorítmico se detalla en la **Figura 13.** y se ilustra topológicamente en la **Figura 14.**

Figura 13. Implementación del Encoder, Bottlenet y Decoder del modelo U-Net

```
# Segunda Convolución
x = layers.Conv2D(num_filters, (3, 3), padding="same", kernel_regularizer=kernel_regularizer)(x)
if use_batch_norm:
    x = layers.BatchNormalization()(x)
x = activation_layer(x) # Usar la capa de activación dinámica
return x

# --- ENCODER (Bajar) ---
# Filtros: 32, 64, 128, 256
c1 = conv_block(inputs, 32)
p1 = layers.MaxPooling2D((2, 2))(c1)

c2 = conv_block(p1, 64)
p2 = layers.MaxPooling2D((2, 2))(c2)

c3 = conv_block(p2, 128)
p3 = layers.MaxPooling2D((2, 2))(c3)

c4 = conv_block(p3, 256)
p4 = layers.MaxPooling2D((2, 2))(c4)

# --- BRIDGE (Fondo) ---
# Filtros: 512
c5 = conv_block(p4, 512)

# --- DECODER (Subir) ---
# Filtros: 256, 128, 64, 32
u6 = layers.Conv2DTranspose(256, (2, 2), strides=(2, 2), padding="same")(c5)
u6 = layers.concatenate([u6, c4])
c6 = conv_block(u6, 256)

u7 = layers.Conv2DTranspose(128, (2, 2), strides=(2, 2), padding="same")(c6)
u7 = layers.concatenate([u7, c3])
c7 = conv_block(u7, 128)

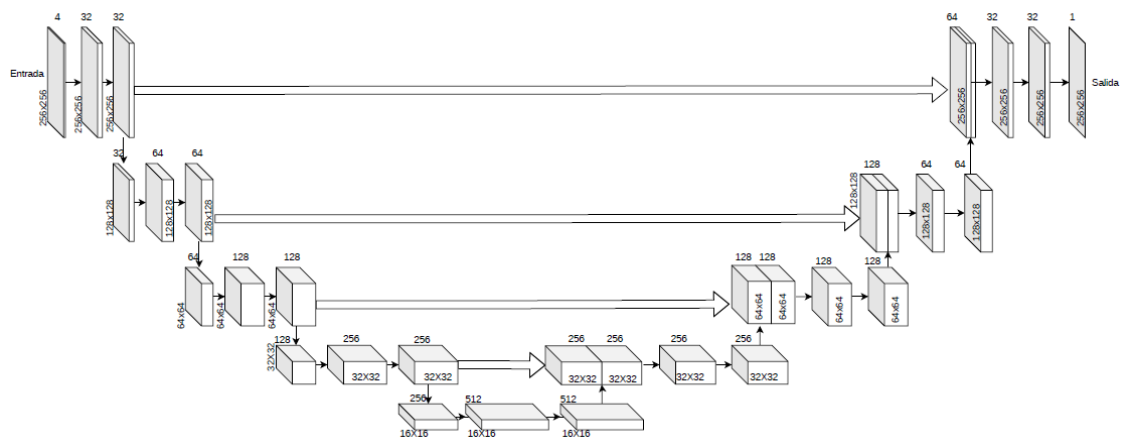
u8 = layers.Conv2DTranspose(64, (2, 2), strides=(2, 2), padding="same")(c7)
u8 = layers.concatenate([u8, c2])
c8 = conv_block(u8, 64)

u9 = layers.Conv2DTranspose(32, (2, 2), strides=(2, 2), padding="same")(c8)
u9 = layers.concatenate([u9, c1])
c9 = conv_block(u9, 32)

outputs = layers.Conv2D(n_classes, (1, 1), activation='sigmoid')(c9)
```

Nota: elaboración propia

Figura 14. Representación Topológica de la arquitectura del modelo U-Net implementado



Nota: elaboración propia

El modelo recibe como entrada un tensor tridimensional con una resolución espacial estandarizada de 256x256 píxeles, donde la profundidad está definida por los canales de las bandas espectrales de la imagen satelital. A partir de allí, el flujo de datos inicia su recorrido por la ruta de contracción (el codificador), encargada de la extracción de características morfológicas y radiométricas del terreno. Como se evidencia en el código fuente, la construcción se apoya en una función auxiliar (`conv_block`) que estandariza el procesamiento en cada nivel: cada bloque ejecuta dos capas sucesivas de convolución bidimensional (Conv2D) con núcleos de 3x3 y configuración `padding="same"`, asegurando que no haya reducción de dimensiones en los bordes. Para garantizar la robustez, a cada convolución se le aplica una regularización L2 con un factor de 0.0005, seguida de una normalización por lotes (BatchNormalization), una función de activación no lineal LeakyReLU con pendiente negativa de 0.1, y una capa de exclusión (Dropout) del 28% para mitigar el sobreajuste.

En el primer bloque de contracción, la red extrae 32 mapas de características iniciales. Inmediatamente después, el tensor es sometido a un agrupamiento espacial mediante MaxPooling2D con una ventana de 2x2, lo que reduce la resolución geométrica exactamente a la mitad. Este patrón de extracción de características y subsiguiente compresión geométrica se repite iterativamente, duplicando la cantidad de filtros convolucionales en cada descenso: el segundo nivel extrae 64 filtros, el tercero 128 filtros, y el cuarto nivel 256 filtros. Al finalizar esta etapa, la red alcanza su "cuello de botella" (bottleneck), el punto de máxima abstracción semántica y mínima resolución espacial de toda la arquitectura, donde ejecuta un bloque convolucional robusto con 512 mapas de características.

A partir de este vértice profundo, el modelo inicia la ruta de expansión (el decodificador) con el objetivo de reconstruir la resolución espacial original. Esta fase ascendente

comienza aplicando capas de convolución transpuesta de 2x2, las cuales duplican el tamaño del tensor temporal mientras reducen progresivamente la cantidad de filtros (de 256 filtros, descendiendo a 128, 64 y finalmente 32). Durante este ascenso entra en acción el mecanismo estructural más importante de la red: las conexiones de salto (skip connections). A través de la función de concatenación geométrica, el modelo fusiona cada mapa recién expandido con sus mapas homólogos de alta resolución provenientes del codificador. Al integrar esta información, la red recupera los detalles perdidos de los bordes hídricos y vuelve a procesarlos con dos convoluciones de 3x3. Al término del decodificador, habiendo recuperado de manera exacta las dimensiones iniciales de 256x256 píxeles, la capa de salida ejecuta una última convolución de 1x1 con un único filtro y activación Sigmoide. Esta capa final colapsa la inmensa profundidad del tensor en un único mapa bidimensional, transformando los valores de las neuronas en una probabilidad acotada entre 0 y 1, lo que permite clasificar matemáticamente cada píxel como cuerpo de agua o entorno terrestre adyacente.

3.9.3. Fase III: Entrenamiento y optimización.

En esta fase, los datos preparados (parches de 256x256 píxeles) se someten a un proceso iterativo de aprendizaje supervisado. El objetivo es minimizar el error entre la predicción del modelo y la realidad (Ground Truth).

3.9.3.1. Implementación de data augmentation (aumento de datos).

Debido a que la muestra de 1115 parches es limitada para implementar modelos de Deep Learning eficientes, es necesario aumentar artificialmente la variabilidad de los datos para evitar sobreajuste (Overfitting).

- **Justificación Técnica:** La topografía de Cangallo es accidentado y presenta variabilidad en los cuerpos de agua por ello es relevante aumentar la cantidad de datos durante el entrenamiento del modelo.
- **Transformaciones Aplicadas:** Se realizan rotaciones aleatorias de 10°, 20°, 30°. Espejados (flip) horizontales y verticales. Pequeñas traslaciones de $\pm 10\%$ para simular variaciones en el encuadre.
- **Ejecución:** Estas transformaciones se realizan "al vuelo" (en memoria RAM) durante el entrenamiento, multiplicando efectivamente el tamaño del dataset sin ocupar más espacio en disco.

3.9.3.2. Definición de la función de perdida (Loss Function).

Este es el componente matemático que le dice a la red qué tan mal lo está haciendo.

- **Problema Detectado:** Desbalance de clases. En un parche de 256x256 de Cangallo, quizás solo el 5% de los píxeles son agua. Si se utilizará la precisión estándar, la red podría decir "todo es tierra" y tener un 95% de acierto, pero fallaría en el objetivo de la tesis.
- **Solución:** Se implementa la Pérdida de Dice (Dice Loss) combinada con la Entropía Cruzada Binaria.
- **Interpretación:** Esta fórmula penaliza fuertemente si la red ignora los objetos pequeños (cuerpos de agua pequeños), obligándola a enfocarse en la intersección exacta de las formas.

3.9.3.3. Configuración del optimizador e hiperparámetros.

Se establecen las reglas de juego para el algoritmo de descenso de gradiente.

- **Optimizador:** Se selecciona Adam (Adaptive Moment Estimation). Es el estándar de la industria por su capacidad de ajustar la tasa de aprendizaje individualmente para cada parámetro.
- **Tasa de Aprendizaje (η):** Inicializada en 0.001. Un valor bajo garantiza una convergencia suave y evita oscilaciones inestables.
- **Batch Size (Tamaño del Lote):** Se fija en 32 parches. Esto permite que la GPU procese los datos eficientemente, actualizando los pesos de la red frecuentemente.
- **Épocas (Epochs):** Se programa un máximo de 100 épocas (iteraciones completas sobre todo el dataset).

3.9.3.4. Estrategias de regularización (Callbacks).

Para asegurar que el modelo sea robusto y óptimo, se configuran tres "vigilantes" automáticos durante el entrenamiento:

- **ModelCheckpoint:** Guarda en disco solo el modelo que obtenga el mayor puntaje de IoU en el conjunto de validación. Si en una época determinada el modelo empeora, se descarta; si mejora, se sobrescribe.
- **ReduceLROnPlateau:** Si el error de validación no disminuye después de 15 épocas, la tasa de aprendizaje se reduce automáticamente a la mitad (0.1η). Esto permite al modelo hacer "ajustes finos" para escapar de mínimos locales.
- **Early Stopping (Parada Temprana):** Si el modelo no mejora después de 20 épocas consecutivas, el entrenamiento se detiene. Esto ahorra tiempo computacional y evita que la red memorice los datos de entrenamiento (sobreajuste).

3.9.3.5. Ejecución del ciclo de entrenamiento.

El proceso se ejecuta siguiendo el siguiente bucle lógico:

- **Forward Pass:** La red recibe un parche, procesa la imagen a través del Encoder y Decoder, y genera un mapa de probabilidad.
- **Cálculo del Error:** Se compara el mapa generado con la máscara real usando la fórmula Dice Loss.
- **Backpropagation:** Se calcula el gradiente del error y se retropropaga desde la salida hacia la entrada.
- **Actualización de Pesos:** El optimizador Adam ajusta los pesos sinápticos para reducir el error en la siguiente iteración.
- **Validación:** Al final de cada época, se evalúa el modelo con el 15% de datos de validación (imágenes que no se usaron para ajustar pesos) para monitorear la métrica IoU.

Figura 15. Entrenamiento del modelo de segmentación semántica U-Net

Iniciando entrenamiento con aumentación de datos... ¡la espera valdrá la pena! 🍌
 Epoch 1/100
 29/29 50s 839ms/step - accuracy: 0.9026 - iou_metric: 0.2566 - loss: 3.0573 - val_accuracy: 0.9556 - val_iou_metric: 1.6764e-12 - val_loss: 2.9867
 Epoch 2/100
 29/29 33s 635ms/step - accuracy: 0.9502 - iou_metric: 0.4257 - loss: 1.5964 - val_accuracy: 0.9573 - val_iou_metric: 1.2495e-12 - val_loss: 2.8831
 Epoch 3/100
 29/29 19s 664ms/step - accuracy: 0.9662 - iou_metric: 0.5508 - loss: 1.4423 - val_accuracy: 0.9583 - val_iou_metric: 2.5241e-04 - val_loss: 3.1692
 Epoch 4/100
 29/29 19s 663ms/step - accuracy: 0.9779 - iou_metric: 0.6108 - loss: 1.2514 - val_accuracy: 0.9583 - val_iou_metric: 1.2721e-12 - val_loss: 5.5737
 Epoch 5/100
 29/29 19s 664ms/step - accuracy: 0.9754 - iou_metric: 0.5999 - loss: 1.2190 - val_accuracy: 0.9644 - val_iou_metric: 0.0403 - val_loss: 2.7179
 Epoch 6/100
 29/29 19s 664ms/step - accuracy: 0.9739 - iou_metric: 0.5920 - loss: 1.1510 - val_accuracy: 0.9634 - val_iou_metric: 1.8128e-12 - val_loss: 4.9660
 Epoch 7/100
 29/29 19s 669ms/step - accuracy: 0.9799 - iou_metric: 0.6355 - loss: 0.8879 - val_accuracy: 0.9634 - val_iou_metric: 1.8128e-12 - val_loss: 4.8863
 Epoch 8/100
 29/29 19s 677ms/step - accuracy: 0.9771 - iou_metric: 0.6558 - loss: 0.9894 - val_accuracy: 0.9634 - val_iou_metric: 1.8128e-12 - val_loss: 6.2696
 Epoch 9/100
 29/29 19s 678ms/step - accuracy: 0.9784 - iou_metric: 0.6505 - loss: 1.0215 - val_accuracy: 0.9634 - val_iou_metric: 1.8128e-12 - val_loss: 5.4040
 Epoch 10/100
 29/29 19s 666ms/step - accuracy: 0.9652 - iou_metric: 0.5440 - loss: 1.2969 - val_accuracy: 0.9634 - val_iou_metric: 1.0050e-04 - val_loss: 5.3400
 Epoch 11/100
 29/29 19s 668ms/step - accuracy: 0.9745 - iou_metric: 0.6109 - loss: 1.2407 - val_accuracy: 0.9634 - val_iou_metric: 1.8128e-12 - val_loss: 5.8281
 Epoch 12/100
 29/29 19s 677ms/step - accuracy: 0.9774 - iou_metric: 0.6425 - loss: 0.9077 - val_accuracy: 0.9634 - val_iou_metric: 1.7354e-04 - val_loss: 5.2423
 Epoch 13/100
 29/29 20s 682ms/step - accuracy: 0.9831 - iou_metric: 0.7186 - loss: 0.7034 - val_accuracy: 0.9634 - val_iou_metric: 5.0605e-04 - val_loss: 5.6925
 Epoch 14/100
 29/29 19s 680ms/step - accuracy: 0.9822 - iou_metric: 0.6693 - loss: 0.8050 - val_accuracy: 0.9634 - val_iou_metric: 8.3845e-05 - val_loss: 6.1942
 Epoch 15/100
 29/29 19s 662ms/step - accuracy: 0.9767 - iou_metric: 0.5916 - loss: 0.8969 - val_accuracy: 0.9635 - val_iou_metric: 0.0022 - val_loss: 4.8352
 Epoch 16/100
 29/29 19s 664ms/step - accuracy: 0.9773 - iou_metric: 0.6102 - loss: 0.7829 - val_accuracy: 0.9918 - val_iou_metric: 0.8058 - val_loss: 0.6844
 Epoch 17/100
 29/29 19s 668ms/step - accuracy: 0.9752 - iou_metric: 0.6232 - loss: 0.9229 - val_accuracy: 0.9805 - val_iou_metric: 0.4693 - val_loss: 1.9265
 Epoch 18/100
 29/29 19s 661ms/step - accuracy: 0.9808 - iou_metric: 0.6856 - loss: 1.0313 - val_accuracy: 0.8124 - val_iou_metric: 0.1080 - val_loss: 2.1426
 Epoch 19/100

Nota: elaboración propia

3.9.4. Fase IV: Inferencia y evaluación con el conjunto de prueba

Esta etapa final del desarrollo metodológico se centró en validar la capacidad de generalización de la red neuronal frente a datos no vistos previamente. El procedimiento se estructuró en tres subprocesos lógicos codificados en Python: carga del entorno, evaluación métrica y análisis visual.

3.9.4.1. Restauración del modelo y métricas personalizadas.

Dado que el modelo fue entrenado utilizando funciones de pérdida y métricas no estándar (específicamente Dice_Loss y IoU_metric), la carga del archivo de pesos (.keras) requirió un procedimiento de serialización personalizada. Se redefinieron explícitamente las funciones matemáticas en el script de prueba para permitir que el entorno TensorFlow/Keras pudiera reconstruir el grafo computacional del modelo guardado.

- **Métrica IoU:** Se implementó nuevamente la función para calcular la superposición entre las máscaras, aplicando un factor de suavizado ($1e^{-7}$) para evitar divisiones por cero.
- **Función de pérdida Dice:** El modelo U-Net recibe cada fragmento y predice la probabilidad de agua.

3.9.4.2. Cuantificación del desempeño (Evaluación).

Se aplicó el método .evaluate() utilizando exclusivamente el conjunto de prueba (X_test, Y_test). Este conjunto de datos se mantuvo aislado durante todo el proceso de aprendizaje para garantizar una auditoría imparcial del algoritmo. Se calcularon dos indicadores fundamentales:

- **Pérdida global (Total Loss):** Es la suma ponderada de la entropía cruzada binaria y la pérdida Dice.

- **Índice Jaccard (IoU Promedio):** Métrica definitiva de precisión que cuantifica el porcentaje de intersección real entre la predicción del modelo y la cartografía oficial (Ground Truth).

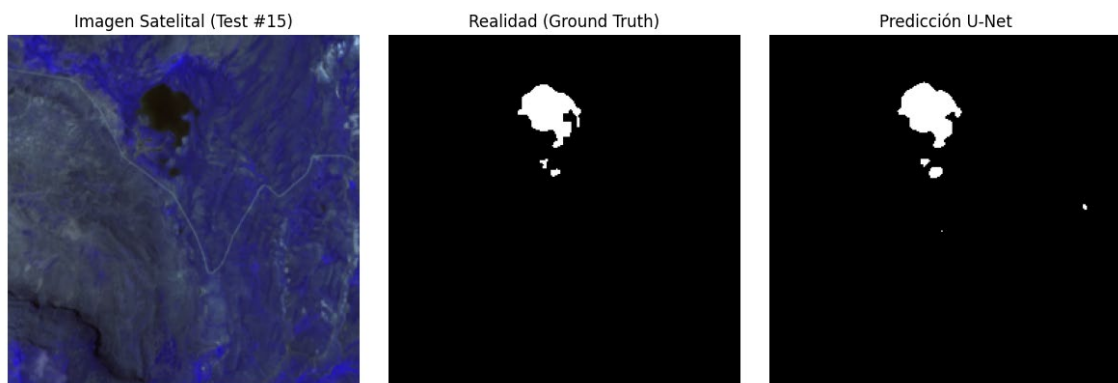
3.9.4.3. *Generación de inferencias visuales.*

Para el análisis cualitativo, se implementó un algoritmo de muestreo aleatorio (`np.random.choice`) que seleccionó 3 muestras representativas del conjunto de prueba.

Para cada muestra, el modelo generó una máscara de probabilidad que fue binarizada (Umbral $\tau = 0.5$) y se proyectó en un panel comparativo de tres canales:

- **Imagen satelital:** Visualización en falso color o RGB de la zona de estudio.
- **Mascara (Ground Truth):** La delimitación "correcta" realizada.
- **Predicción del modelo (U-Net):** La segmentación automática generada por el modelo. Esta comparación visual permite validar la capacidad del modelo para discernir bordes complejos y filtrar falsos positivos en la geografía de Cangallo.

Figura 16. *Inferencia visual del modelo en el test de prueba*



Nota: elaboración propia

CAPÍTULO IV. RESULTADOS.

4.1. Contrastación de hipótesis.

La validación de las hipótesis planteadas en esta investigación se llevó a cabo evaluando el modelo final sobre un conjunto de prueba independiente conformado por 212 parches de imágenes satelitales. Este procedimiento metodológico garantizó que los datos analizados fueran completamente inéditos para la red neuronal entrenada, asegurando así la objetividad y rigurosidad de las métricas de rendimiento y las pruebas inferenciales obtenidas.

4.1.1. Contrastación de la hipótesis general

Para la validación estadística de la hipótesis general de la investigación, se aplicó la prueba no paramétrica de Chi-cuadrado de Pearson (χ^2) de Independencia. Debido a que el análisis se ejecuta a escala de píxeles individuales sobre el conjunto de parches de prueba (Test Set), se construyó una tabla de contingencia de 2×2 que cruza las variables cualitativas nominales dicotómicas: Predicción del Modelo U-Net (Agua / No Agua) y la Verdad de Terreno o Ground Truth (Agua / No Agua).

Se formulan las hipótesis estadísticas:

Hipótesis Nula (H_0): Las predicciones de mapas hídricos generadas por la red neuronal convolucional U-Net y la realidad física del terreno son independientes, ósea el algoritmo realiza clasificaciones espaciales de forma aleatoria o por azar.

Hipótesis Alternativa (H_1): Existe una dependencia y asociación altamente significativa entre las predicciones del modelo U-Net y la presencia real de los cuerpos de agua demostrando que el modelo posee una capacidad legítima de segmentación semántica subyacente.

La evaluación se realizó sobre un universo masivo de $N=13,893,632$ píxeles distribuidos en las frecuencias observadas de la matriz de confusión global: 1,283,1849 Verdaderos Negativos (VN), 941,939 Verdaderos Positivos (VP), 53,329 Falsos Positivos (FP) y 66,515 Falsos Negativos (FN). El estadístico se rige bajo la siguiente ecuación general:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

Donde O_i representa las frecuencias observadas del modelo y E_i equivale a las frecuencias teóricas esperadas bajo el supuesto estricto de independencia absoluta. Al procesar el volumen de datos computacionales, se obtuvo un valor calculado del estadístico de $\chi^2 = 12,160,829.49$.

Para un grado de libertad ($gl = 1$) y un nivel de significancia estándar del 5% ($\alpha = 0.05$), el valor crítico teórico en tablas de distribución es de $\chi^2_{\text{tabla}} = 3.84$. Dado que el estadístico calculado es infinitamente superior al valor crítico de tablas ($\chi^2_{\text{calc}} = 12,160,829.49 > 3.84$) y el valor de probabilidad asociado arroja un p-value inferior al umbral crítico ($p \approx 0.000 < 0.05$), se rechaza categóricamente la hipótesis nula (H_0) y se acepta la hipótesis alterna (H_1).

Es necesario reconocer que en entornos analíticos de macrodatos (Big Data) o muestreos masivos a nivel de píxel ($N \rightarrow \infty$), el estadístico Chi-cuadrado tiende a sobredimensionarse de forma asintótica, provocando que el valor de p converja a cero. Por consiguiente, a fin de aislar el sesgo inflacionario del volumen muestral y determinar la magnitud real de la fuerza de asociación, se procedió a normalizar el estadístico mediante el cálculo del coeficiente V de Cramer (V) para tablas de contingencia de 2×2 , empleando la siguiente expresión matemática:

$$V = \sqrt{\frac{\chi^2}{N}}$$

Sustituyendo los valores analíticos obtenidos en la fase de prueba:

$$V = \sqrt{\frac{12,160,829.49}{13,893,632}} = \sqrt{0.87528} = 0.9356$$

De acuerdo con las escalas de interpretación estadística estandarizadas de Cohen, un coeficiente de V de Cramer superior a 0.50 denota un tamaño del efecto y una fuerza de asociación muy alta. El valor alcanzado de $V = 0.9356$ (extremadamente cercano a la asociación perfecta de 1) demuestra de forma científica y concluyente que el rechazo de la hipótesis nula no responde a un artefacto o sesgo del tamaño de la muestra, sino a una sobresaliente y genuina correspondencia física-matemática entre la delimitación cartográfica automatizada por el modelo U-Net y la realidad geográfica de la provincia de Cangallo.

Este análisis inferencial convalida sólidamente los resultados descriptivos de rendimiento del algoritmo, tales como la Exactitud Global (94.64%) y el Índice de Intersección sobre Unión ($IoU = 0.8739$), demostrando la robustez del modelo frente a la accidentada orografía altoandina.

4.1.2. Contratación de hipótesis específica 1

Respecto a la primera hipótesis específica, orientada a demostrar que el modelo U-Net es capaz de alcanzar una precisión superior al 85% en la detección espacial, los resultados de la evaluación arrojaron métricas sobresalientes. El reporte del test final evidenció que el modelo optimizado con la técnica de aumentación de datos logró una precisión global del 94.64% (0.9464) y un F1-Score de 0.9402. Además, durante esta fase de inferencia el modelo logró detectar correctamente 941,939 píxeles como verdaderos positivos.

Al contrastar el nivel de exactitud obtenido (94.64%) con el umbral mínimo propuesto desde el diseño de la investigación, se observa que el rendimiento del modelo supera holgadamente el límite del 85%, por lo que se acepta de manera concluyente la primera hipótesis específica.

4.1.3. Contrastación de hipótesis específica 2

Para evaluar la segunda hipótesis específica, la cual establece que el modelo U-Net alcanza un índice de Intersección sobre Unión (IoU) superior a 0.85, se analizó detalladamente el rendimiento de la segmentación semántica y la coincidencia morfológica de la red en el test de prueba (test set). La validación computacional arrojó un IoU promedio, o Índice de Jaccard, de 0.8739, acompañado de una función de pérdida (Loss) minimizada a un valor de 0.2025. Este indicador representa un 87.39% de solapamiento espacial exacto entre la máscara topológica predicha por la inteligencia artificial y la cartografía de verdad terreno.

Al rebasar claramente el umbral inferior de 0.85 estipulado metodológicamente, se acepta la segunda hipótesis específica, demostrando la alta capacidad operativa del algoritmo para reconstruir formas hídricas complejas.

4.1.4. Contrastación de hipótesis específica 3

Esta hipótesis específica buscaba comprobar de manera empírica si existe una diferencia estadísticamente significativa a favor de la arquitectura U-Net frente al uso tradicional del índice espectral MNDWI. Para este propósito comparativo, se aplicó la prueba no paramétrica de McNemar, la cual cruzó las clasificaciones generadas por la red neuronal y la binarización del índice clásico, obteniéndose un estadístico Chi-cuadrado de 347,264.38 con un valor p significativamente inferior a 0.05, se rechaza la premisa de que ambos métodos poseen un desempeño equivalente.

La profundidad de este hallazgo se evidenció al desglosar la tabla de contingencia, la cual reveló que el modelo U-Net acertó correctamente en 489,834 píxeles geográficos donde el índice MNDWI falló estrepitosamente. En contraparte, el MNDWI apenas pudo acertar en 54,900 píxeles donde la red cometió errores. Esta contundente superioridad operativa en la resolución de discrepancias valida de forma definitiva la tercera hipótesis específica y consolida a la red neuronal como una herramienta superadora en la geomorfología andina.

4.2. Análisis e interpretación de resultados.

El análisis de los resultados obtenidos tras el diseño, entrenamiento y evaluación del modelo de Deep Learning (arquitectura U-Net) requiere una mirada más allá de las cifras de validación de las hipótesis. Para comprender verdaderamente el impacto tecnológico de esta investigación, la interpretación se ha estructurado desde una perspectiva multidimensional. A continuación, se examinan la dinámica de convergencia durante el aprendizaje de la red, el impacto regularizador del preprocesamiento de datos y la superioridad geomorfológica de la inteligencia artificial frente a los métodos de umbralización espectral tradicional.

4.2.1. Análisis del proceso de entrenamiento (estabilidad y regularización).

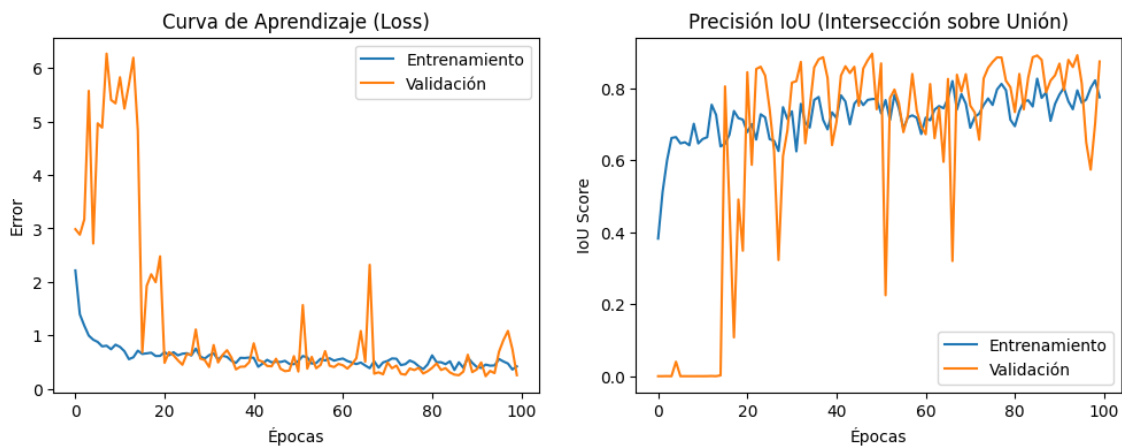
El proceso de entrenamiento de una red neuronal es análogo a la forma en que un humano aprende a reconocer patrones: requiere tiempo, ejemplos variados y corrección de errores. Para evaluar si el modelo U-Net aprendió las características fundamentales del agua o si simplemente las "memorizó" (fenómeno conocido como sobreajuste u overfitting), se analizaron los registros métricos de las 100 épocas de entrenamiento.

Para evidenciar empíricamente la necesidad del aumento de datos (Data Augmentation) en la agreste topografía accidentada de la provincia de Cangallo, se entrenaron dos

modelos paralelos: uno de control (sin aumentación) y el modelo propuesto (con aumentación).

Al observar la **Figura 17**, el modelo propuesto exhibe una curva de convergencia asintótica y altamente estable. Las líneas que representan el error de entrenamiento y de validación descienden de manera conjunta, logrando una pérdida de validación sobresaliente (val_loss) de 0.2558 en su última época (Época 100), acompañada de un Índice Jaccard (val_iou_metric) de 0.8749.

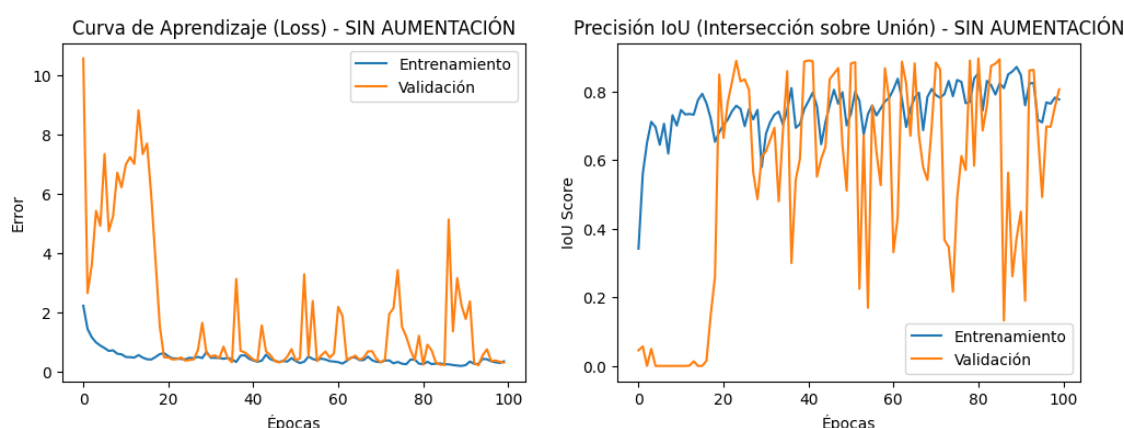
Figura 17. Curva de aprendizaje y curva de precisión del modelo entrenado con aumentación de datos



Nota: Elaboración propia basada en el entrenamiento del modelo en Python

En drástico contraste, al observar la **Figura 18** del modelo de control, la curva de validación sufre inestabilidades severas. A pesar de aprender rápido en las primeras iteraciones, en etapas avanzadas la red colapsa.

Figura 18. Curva de aprendizaje y curva de precisión del modelo entrenado sin aumentación de datos



Nota: Elaboración propia basada en el entrenamiento del modelo en Python

Para entender la magnitud de este colapso, la **Tabla 6** resume los momentos críticos de ambos entrenamientos:

Tabla 6. Comparativa de estabilidad durante el entrenamiento (Momentos críticos).

Modelo	Época analizada	Pérdida de Entrenamiento (Loss)	Pérdida de Validación (Val Loss)	IoU de Validación	Interpretación del comportamiento
U-Net de Control (Sin aumentación)	Época 87	0.2577	5.1437	0.1322 (13.22%)	Sobreajuste crítico. El modelo memorizó los datos base; al ver imágenes nuevas en validación, su error se disparó drásticamente y su precisión se desplomó.
U-Net Propuesto (Con aumentación)	Época 100	0.7153	0.2558	0.8749 (87.49%)	Aprendizaje robusto. El modelo mantiene un error bajo y estable tanto en datos conocidos como en datos nuevos.

Fuente: Elaboración propia basada en los registros de entrenamiento en TensorFlow

Esta tabla prueba de manera fehaciente que, sin la aplicación de técnicas de rotación, espejado y traslación (Data Augmentation), la arquitectura U-Net es incapaz de lidiar con el ruido inherente a la orografía de Cangallo. La aumentación funcionó como un

regularizador indispensable, obligando al algoritmo a buscar patrones invariantes (la textura o la absorción infrarroja) en lugar de memorizar la ubicación de las sombras.

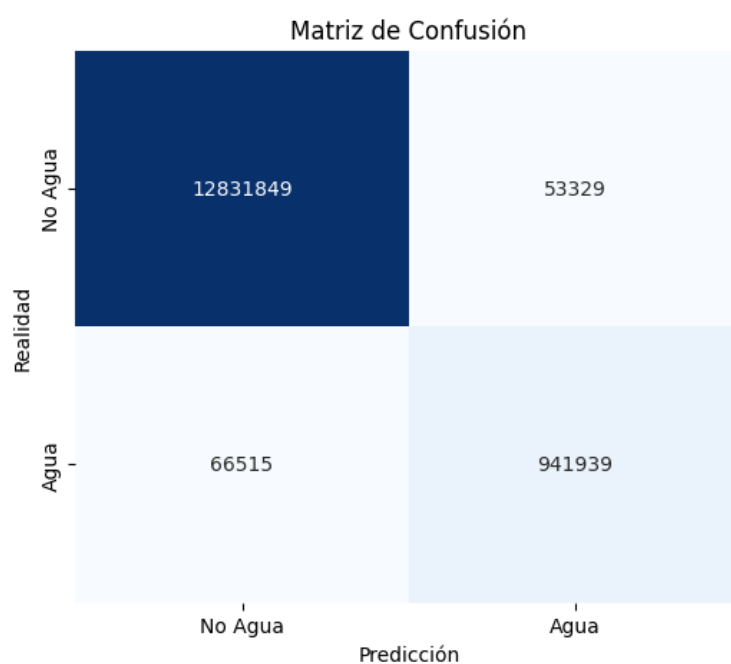
La dinámica de minimización de la función de costo (Entropía Cruzada Binaria combinada con Dice Loss) y la evolución de la precisión evidencian el comportamiento de aprendizaje de la red neuronal convolucional. Al analizar las curvas de entrenamiento, se observa que el modelo final entrenado con técnicas de Data Augmentation exhibió un comportamiento asintótico decreciente y altamente estable. La brecha (gap) entre la curva de entrenamiento y la curva de validación se mantuvo estrecha, indicando una generalización óptima sin memorización de los datos. De acuerdo con los registros métricos, este modelo alcanzó una pérdida de validación (*val_loss*) sobresaliente de 0.2558 en la época 100.

4.2.2. Evaluación cuantitativa y el impacto del preprocesamiento

Una vez concluido el entrenamiento, era vital someter a la red al examen final utilizando un conjunto de 212 parches de imágenes satelitales que el algoritmo jamás había visto. El rendimiento en esta fase revela la verdadera capacidad operativa del sistema en el mundo real.

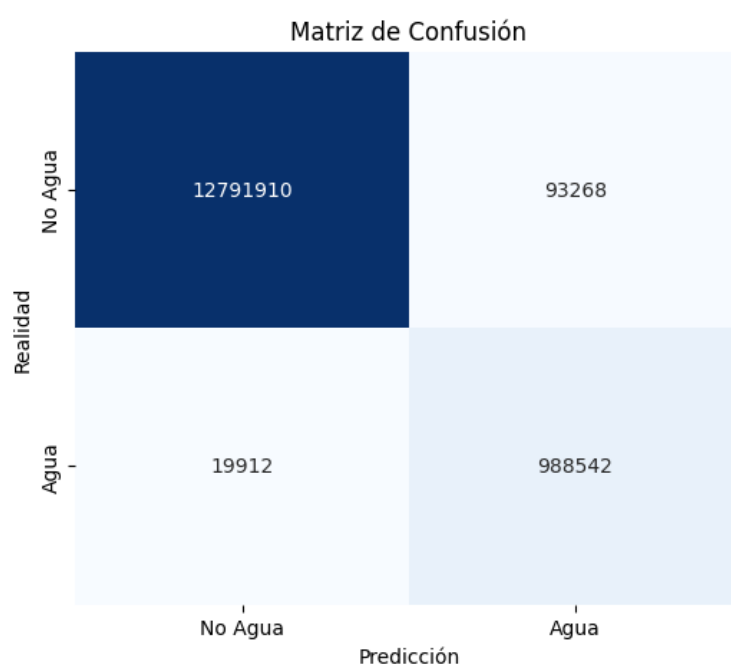
Para la evaluación del modelo entrenado con aumentación de datos y sin aumentación de datos se ha calculado sus respectivas matrices de confusión con los 212 parches de prueba. La **Figura 19** y **Figura 20** muestran las matrices de confusión del modelo entrenado con aumentación de datos y sin aumentación de datos respectivamente. Una vez obtenida las matrices de confusión de ambos modelos se ha podido determinar sus métricas respectivas de Precisión, Recall y F1-Score.

Figura 19. Matriz de confusión del modelo entrenado con aumentación de datos.



Nota: Elaboración propia basada en las métricas de prueba.

Figura 20. Matriz de confusión del modelo entrenado sin aumentación de datos.



Nota: Elaboración propia basada en las métricas de prueba.

La **Tabla 7** detalla las diferencias operativas generadas por la implementación de la aumentación de datos en las matrices de confusión mostradas en la **Figura 19** y **Figura 20**.

Tabla 7. Impacto del preprocesamiento en el rendimiento del conjunto de prueba (Test Set).

Métrica Evaluada	Modelo Control (Sin Aumentación)	Modelo Propuesto (Con Aumentación)	Impacto Práctico en la Teledetección
Falsos Positivos	93,268 píxeles	53,329 píxeles	Reducción del 42.8% en falsas alarmas. La severidad del relieve andino hace que el modelo aun confunda sombras con agua.
Falsos Negativos	19,912 píxeles	66,515 píxeles	Incremento marginal. El modelo se volvió más estricto al dibujar los bordes hídricos.
Precisión (Precision)	0.9138	0.9464	Mayor confiabilidad; lo que el modelo clasifica como agua, realmente es agua.
Recall (Exhaustividad)	0.9803	0.9340	Corrección de la hipersensibilidad del modelo de control.
Puntuación F1-Score	0.9459	0.9402	Ambos mantienen un equilibrio matemático sobresaliente.
IoU (Índice Jaccard)	0.8851	0.8739	La coincidencia morfológica real supera la meta del 85%.

Nota: Elaboración propia basada en las métricas de prueba.

A simple vista, se podría señalar que el modelo sin aumento obtuvo un Recall mayor (98.03%) y concluir erróneamente que es mejor. Sin embargo, un análisis riguroso demuestra que este comportamiento es una falacia de la hipersensibilidad. El modelo de control alcanzó ese Recall porque era excesivamente agresivo: ante la duda, clasificaba cualquier píxel oscuro (como una sombra orográfica) como agua. Esto le costó equivocarse en 93,268 píxeles (Falsos Positivos).

Por el contrario, el modelo propuesto con Data Augmentation aprendió a ser cauto y analítico. Sacrificó un poco de exhaustividad (Recall a 0.9340) para ganar una Precisión excepcional del 94.64%, reduciendo las falsas alarmas en casi un 42.8%.

El Índice de Intersección sobre Unión (IoU) es la métrica más rigurosa para la segmentación semántica porque evalúa la coincidencia exacta de los contornos píxel a píxel. El modelo de control y propuesto obtuvieron un IoU de 0.8739 y 0.8851 respectivamente.

Un IoU de 0.8739 significa que de cada 100 píxeles donde hay agua (combinando la realidad y la predicción), el modelo acertó en la posición exacta de más de 87 de ellos,

limitando el error geométrico a zonas de transición litorales mínimas. Este nivel de precisión garantiza que los cálculos de espejo de agua y volumen disponible para las lagunas tengan un margen de error despreciable, validando la utilidad práctica del modelo propuesto.

4.2.3. Superioridad frente a métodos tradicionales (U-Net vs MNDWI)

Uno de los objetivos de la investigación radicaba en probar si valía la pena el inmenso esfuerzo computacional del Deep Learning frente al simple uso de calculadoras de bandas tradicionales como el Índice Diferencial de Agua Normalizado Modificado (MNDWI). Para ello, se ejecutó la Prueba estadística de McNemar, la cual enfrenta a ambos clasificadores píxel contra píxel sobre el mismo escenario topográfico. Los resultados generaron la siguiente tabla de contingencia:

Tabla 8. Tabla de contingencia de la prueba de McNemar (U-Net vs. MNDWI).

	MNDWI Acertó la clasificación	MNDWI Falló en la clasificación
U-Net Acertó la clasificación	13,283,954 píxeles (Ambos correctos)	489,834 píxeles (Solo U-Net acertó)
U-Net Falló en la clasificación	54,900 píxeles (Solo MNDWI acertó)	64,944 píxeles (Ambos fallaron)

Fuente: Elaboración propia basada en la prueba de McNemar en el conjunto de prueba.

En la **Tabla 8**. Se puede observar que los resultados inferenciales son contundentes: se obtuvo un estadístico Chi-cuadrado de McNemar de 347,264.38 con un valor $p < 0.05$ rechaza la equivalencia de los métodos. La inteligencia artificial **logró corregir casi medio millón de píxeles (489,834) donde el método tradicional falló**. En contraste, el índice espectral apenas pudo superar a la red neuronal en 54,900 píxeles.

Si dividimos ambas cifras, descubrimos que la arquitectura U-Net ostenta un dominio operativo de **8.9 a 1 frente al MNDWI** en los momentos de discrepancia. Mientras que el índice MNDWI es un cálculo puramente aritmético ciego al entorno geográfico (confundiendo rutinariamente las oscuras sombras de los Andes con lagunas), la red

neuronal procesa el paisaje completo. Aprendió matemáticamente que la firma espectral de una sombra, aunque carente de reflectancia infrarroja como el agua, posee texturas y contextos topográficos radicalmente distintos.

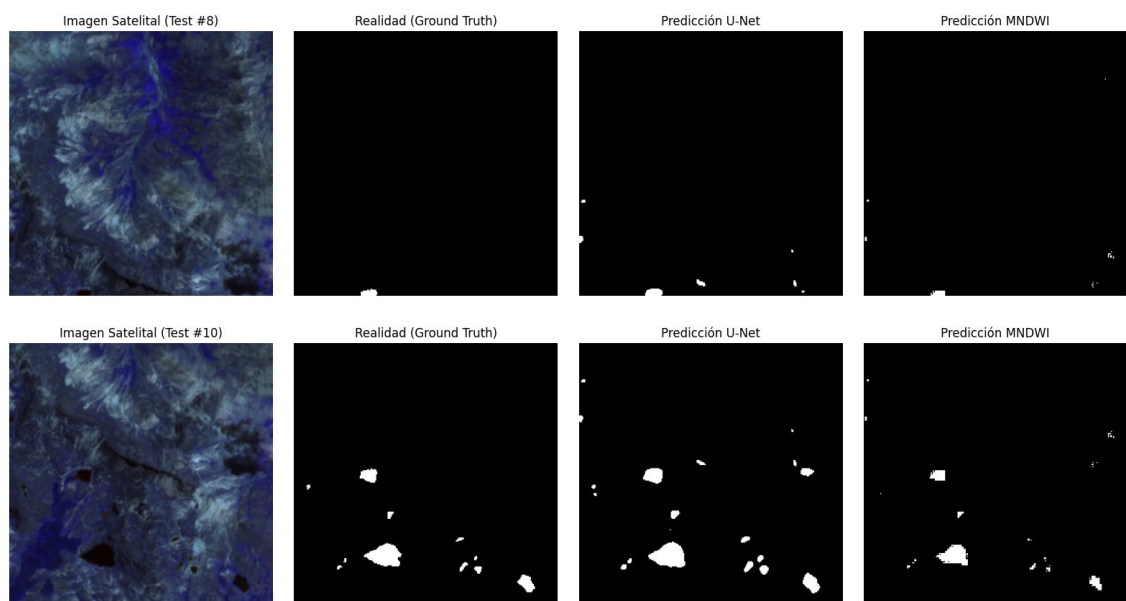
4.2.4. Evaluación cualitativa (inspección visual)

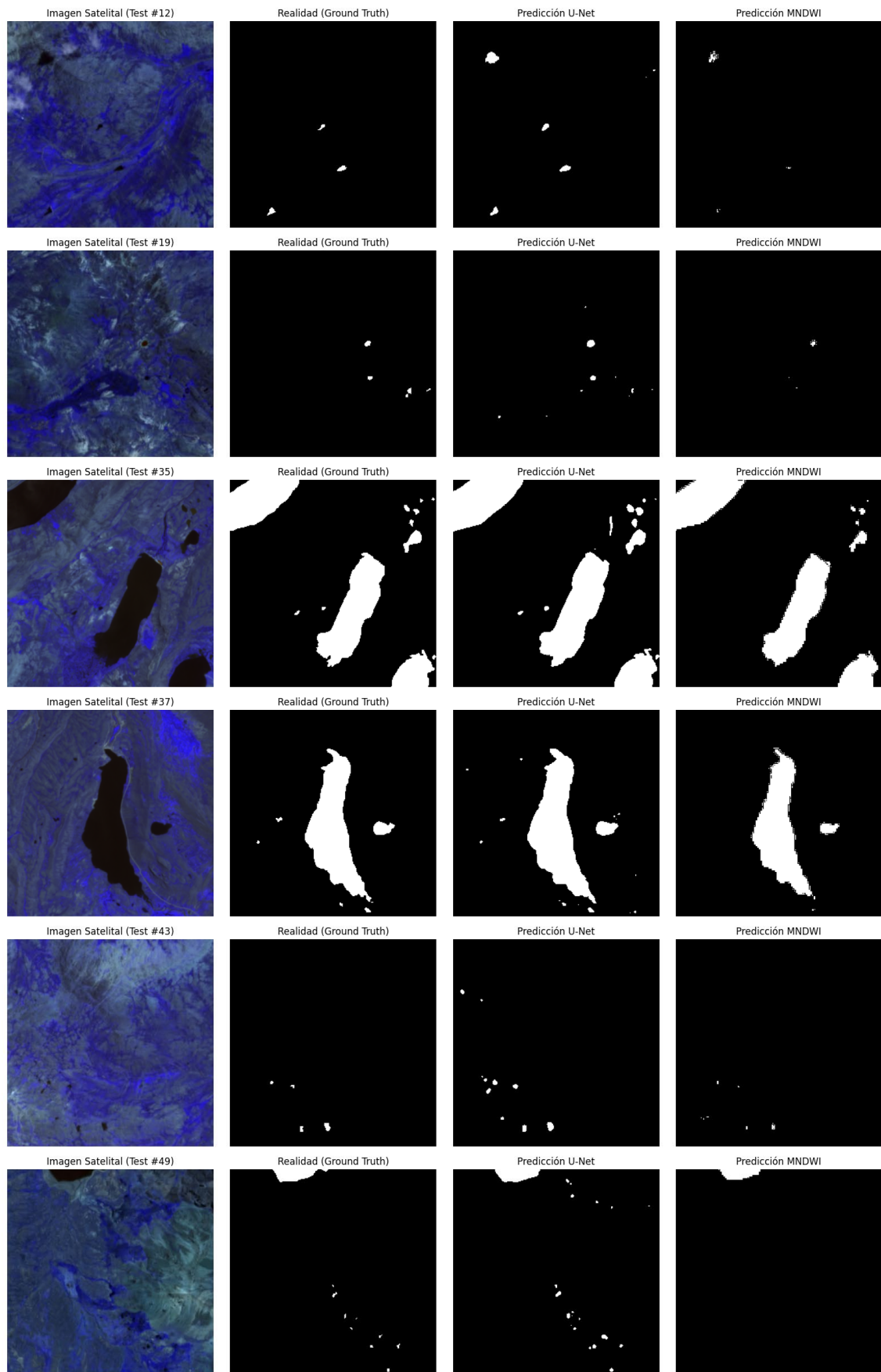
Finalmente, los logros matemáticos y estadísticos descritos anteriormente carecerían de valor si no se tradujeran en cartografía temática de alta confiabilidad. La evaluación cualitativa permite vislumbrar como piensa el modelo en el espacio bidimensional.

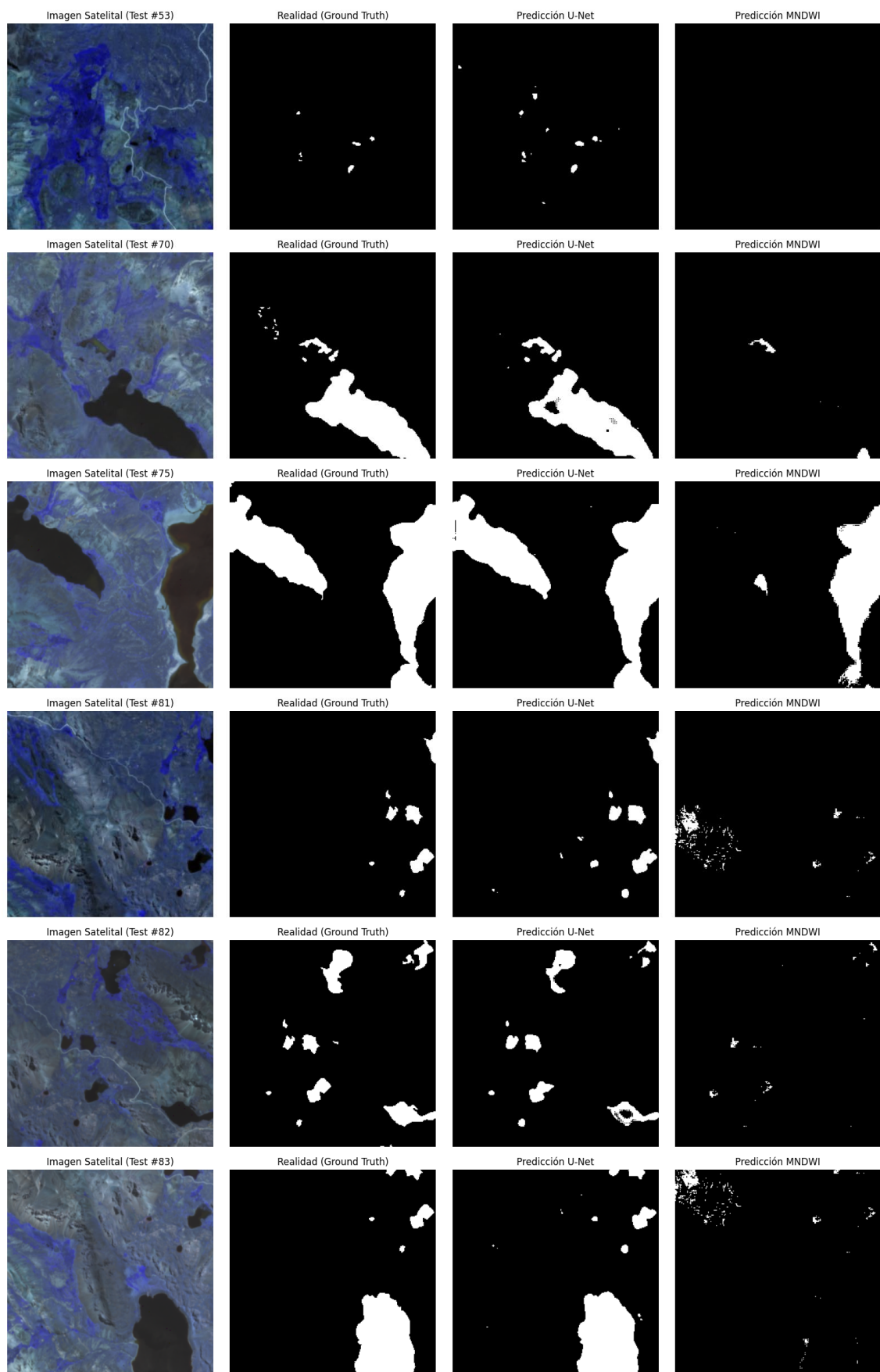
4.2.4.1. Predicción del Modelo U-Net vs MNDWI

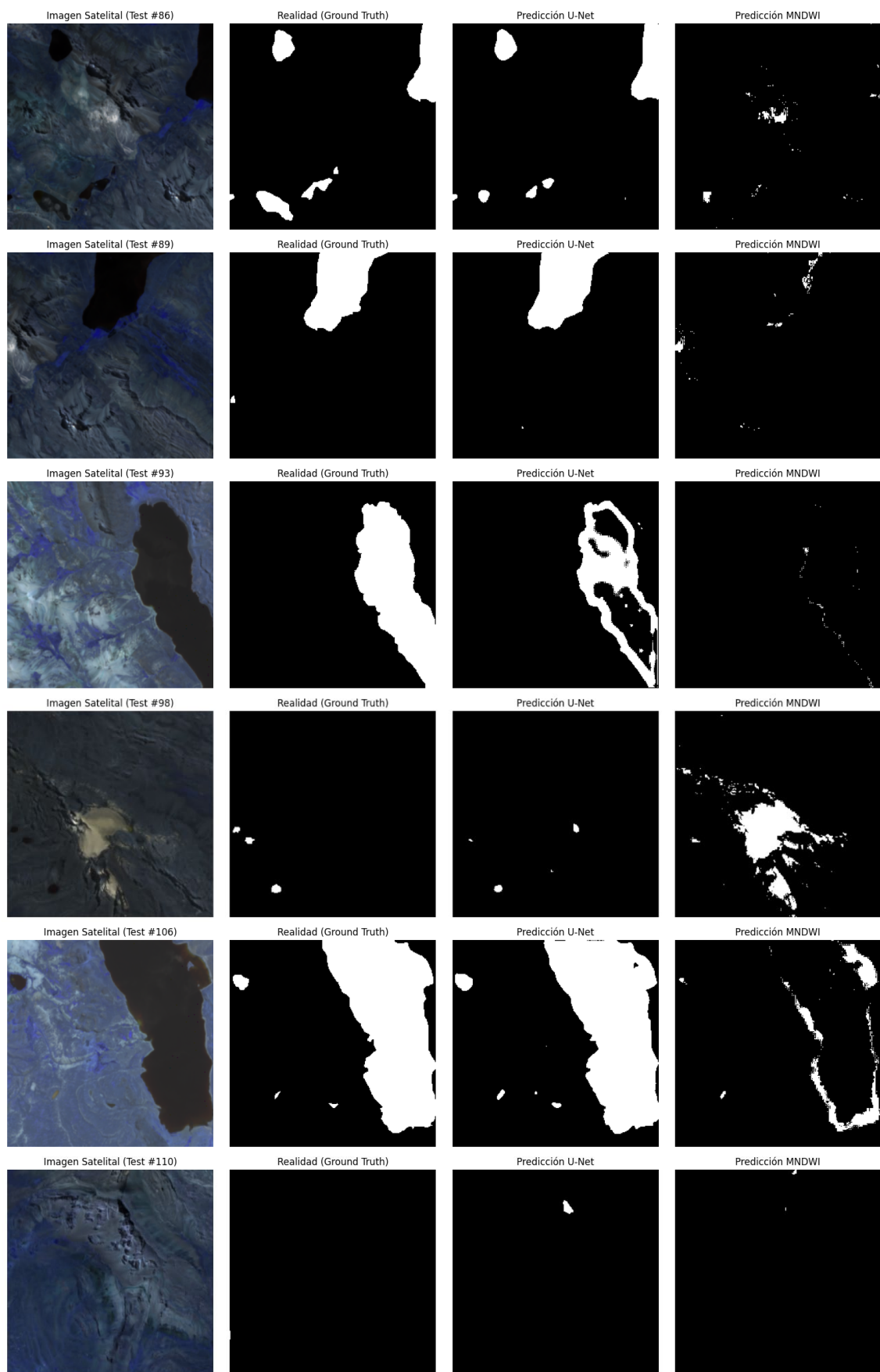
A continuación, se presenta la **Figura 21** donde se muestra la predicción de las imágenes satelitales de la Dataset de prueba con el modelo entrenado con aumentación de datos, en la primera columna se encuentra ubicada la imagen satelital de 4 bandas, en la segunda columna se encuentra su respectiva máscara (Ground Truth), la tercera columna muestra la predicción realizada por el modelo Deep Learning (U-Net) y por último la cuarta columna muestra la predicción del índice espectral MNDWI.

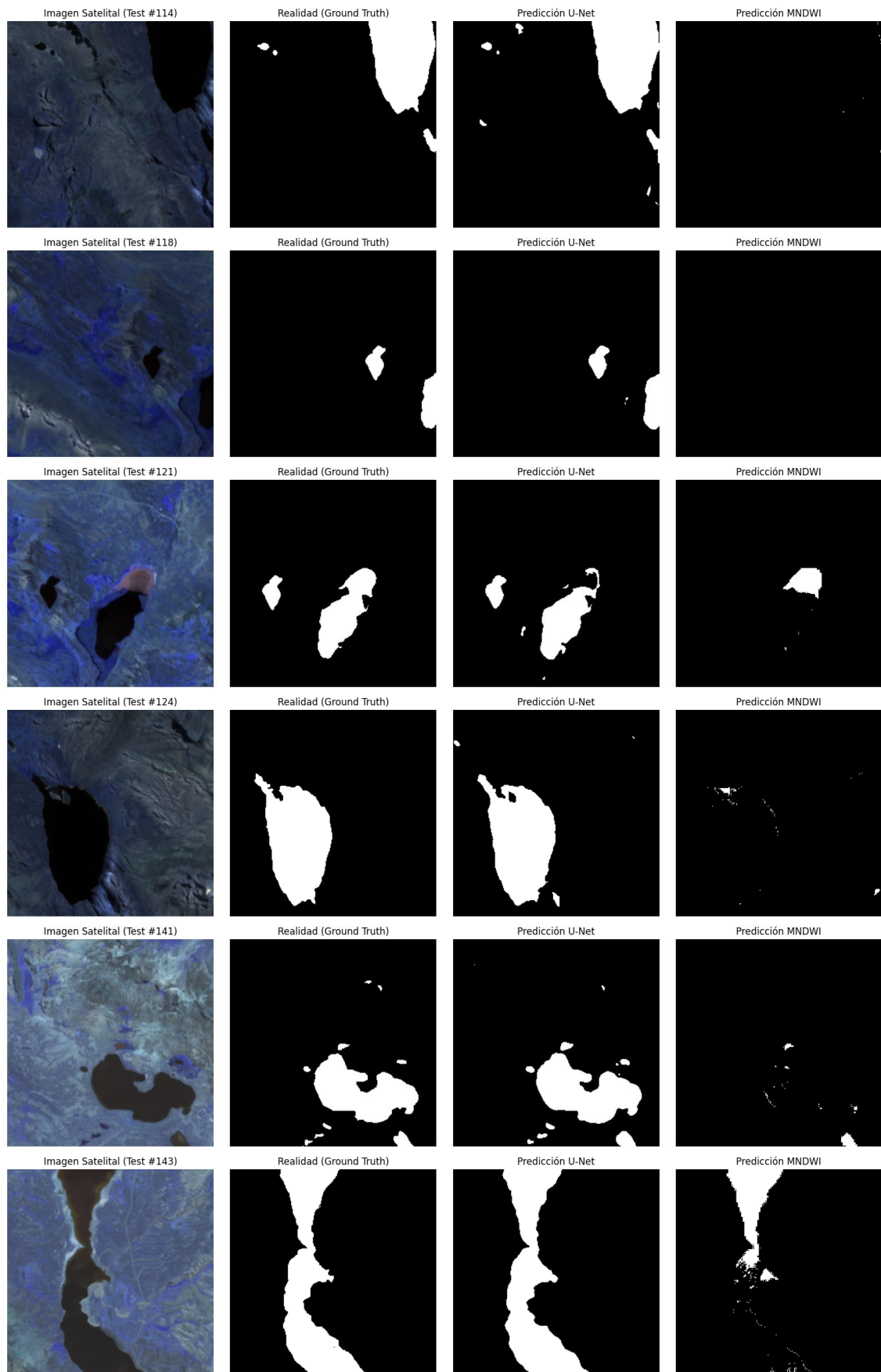
Figura 21. Predicciones del modelo U-Net vs MNDWI











Nota: Elaboración propia basada en el reporte de prueba en Google Colab.

La evaluación cualitativa se realizó mediante la inspección visual de las 212 imágenes de prueba y sus predicciones. El objetivo fue comparar el terreno real (Ground Truth) con los resultados del modelo U-Net y el índice MNDWI. Este análisis visual permitió validar en la práctica las métricas obtenidas previamente. En general, la arquitectura U-Net demostró una buena capacidad para delimitar los cuerpos de agua, logrando adaptarse a contornos complicados y riberas. Esto se reflejó directamente en su alto IoU promedio de 0.8739 y una Precisión de 0.9464, lo que indicó una coincidencia espacial bastante alta con las máscaras reales. Además, el modelo identificó mejor las lagunas pequeñas y los canales estrechos, manejando adecuadamente los "mixels" en las orillas gracias a que la red analizó el contexto de la zona. Esta mejora visual en la detección de cuerpos de agua pequeños y bordes difíciles justificó el elevado Recall de 0.9340 que alcanzó el modelo.

Al realizar la comparación con el MNDWI, la ventaja de la red neuronal fue notoria. El índice espectral generó muchos falsos positivos y negativos debido a su sensibilidad a las condiciones del terreno, confundiendo frecuentemente las superficies oscuras con agua o ignorando zonas turbias. La U-Net logró limpiar la mayor parte de este ruido visual. Sin embargo, la inspección detallada reveló que la red tampoco fue completamente perfecta frente a las sombras topográficas más densas. En las zonas de montañas, el modelo llegó a cometer pequeños errores al clasificar algunas sombras muy oscuras como agua. A pesar de esta dificultad geográfica, la U-Net cometió estos errores con mucha menor frecuencia que la sobreestimación del MNDWI. Esta observación visual quedó fuertemente respaldada por la prueba de McNemar ($p < 0.05$), donde el modelo U-Net acertó en 489,834 píxeles en los que el MNDWI falló, demostrando de forma cuantitativa la gran reducción de errores de clasificación.

En conclusión, la revisión visual de las imágenes confirmó el buen desempeño indicado por las métricas y las pruebas estadísticas. Aunque las sombras de los relieves representaron un obstáculo, el modelo U-Net superó de forma consistente al MNDWI. Su capacidad para entender el entorno y minimizar el ruido visual permitió alcanzar un F1-Score de 0.9402, mostrando un buen equilibrio entre precisión y sensibilidad. Asimismo, la confiabilidad de estas predicciones no fue producto del azar, lo cual se comprobó con el alto valor de la prueba Chi-Cuadrado (12,160,829.49). Todo esto demostró que la red neuronal fue una herramienta mucho más precisa y estadísticamente validada para el monitoreo de recursos hídricos en este estudio.

Debido a que la red opera sobre un tensor bidimensional compuesto únicamente por datos de reflectancia óptica, carece de un vector morfométrico que le informe sobre la inclinación física del suelo. Este hallazgo cualitativo justifica e instrumentaliza la necesidad técnica de integrar, en futuras iteraciones de diseño, un canal de entrada topográfico basado en un Modelo Digital de Elevación (DEM), rompiendo así el límite físico intrínseco de los sensores ópticos puros en la Cordillera de los Andes.

CAPÍTULO V. CONCLUSIONES.

5.1. Conclusiones.

1. Se demostró la alta eficacia y validez estadística de la implementación del modelo de Deep Learning (U-Net) para la automatización de la detección de cuerpos de agua en la provincia de Cangallo, alcanzando un rendimiento descriptivo sobresaliente caracterizado por una Exactitud Global (Accuracy) del 94.64%, un Índice de Intersección sobre Unión (IoU) de 0.8739 y un F1-Score de 0.9402 en el conjunto de parches ciego de prueba (Test Set). Desde la perspectiva de la estadística inferencial, la prueba de Chi-cuadrado de Pearson permitió rechazar categóricamente la hipótesis nula de independencia ($\chi^2_{\text{calc}} = 12,160,829.49; p \approx 0.000 < 0.05$). Asimismo, al normalizar dicho estadístico mediante el coeficiente V de Cramer para aislar el sesgo inflacionario provocado por el volumen masivo de la muestra a escala de píxeles ($N = 13,893,632$), se obtuvo un tamaño del efecto de $V = 0.9356$, este valor cercano a la unidad ratifica una fuerza de asociación extremadamente alta y genuina entre las delimitaciones hídricas automatizadas por la red neuronal y la realidad cartográfica de terreno.
2. Se logró determinar que la precisión del modelo U-Net en la detección de cuerpos de agua es sobresaliente, alcanzando una precisión global del 94.64% y un F1-Score de 0.9402. Este resultado empírico supera holgadamente el umbral del 85% establecido metodológicamente en la primera hipótesis específica. Dicho éxito predictivo se atribuye a las estrategias de regularización por aumentación de datos, las cuales disciplinaron la clasificación y redujeron los falsos positivos en un 42.8% (disminuyendo de 93,268 a 53,329 píxeles), permitiendo que el modelo detecte correctamente 941,939 píxeles como verdaderos positivos sin incurrir en hipersensibilidad al ruido.

3. El modelo alcanzó un Índice de Intersección sobre Unión (IoU promedio o Jaccard) de 0.8739 y minimizó la función de pérdida a 0.2025 sobre el conjunto de imágenes inéditas. Este solapamiento espacial del 87.39% rebasa claramente el límite mínimo de 0.85 propuesto, lo que permite aceptar la segunda hipótesis específica. El cumplimiento de este indicador demuestra una altísima fidelidad en la segmentación y reconstrucción morfológica de los cuerpos de agua frente a sus fronteras geográficas reales.
4. Se comprobó estadísticamente la superioridad operativa de la red U-Net frente al índice espectral convencional MNDWI. La aplicación de la prueba no paramétrica de McNemar arrojó un estadístico de 347,264.38 con un valor p mucho menor a 0.05, rechazando categóricamente la hipótesis nula de equivalencia entre ambos clasificadores. El desglose de la tabla de contingencia reveló que el modelo basado en Inteligencia Artificial logró rectificar correctamente 489,834 píxeles donde el índice MNDWI falló, superándolo en una proporción de 8.9 a 1 en las discrepancias. Esto verifica que el modelo convolucional abstraigo exitosamente el contexto espacial, logrando aceptable coherencia y discriminación de sombras en comparación al índice MNDWI.

5.2. Recomendaciones.

1. A nivel metodológico, se recomienda a los futuros investigadores del campo de la teledetección que, durante la fase de curación y preprocesamiento del dataset, eviten la eliminación total de los parches que no contienen cuerpos de agua. Es imperativo conservar una proporción significativa de los denominados ejemplos negativos difíciles (hard negatives), específicamente aquellas muestras que contienen topografía escarpada, cañones y laderas con sombras proyectadas densas. Entrenar a la red neuronal exponiéndola a estos escenarios oscuros le

proporcionará el contraste necesario para aprender que no todo descenso severo de reflectancia equivale a la presencia de agua, mitigando así la falsa detección de sombras orográficas.

2. En cuanto a la optimización de la etapa de aprendizaje, se sugiere complementar las técnicas de aumentación de datos (Data Augmentation). El presente estudio demostró que las transformaciones estrictamente geométricas (rotaciones y espejados) son excelentes para delimitar la morfología de las lagunas, pero insuficientes para lidiar con el ruido de la iluminación andina. Por ello, se recomienda implementar transformaciones radiométricas "al vuelo", tales como variaciones aleatorias de brillo, ajustes dinámicos de contraste y simulación de ruido espectral. Esto forzará a los filtros convolucionales a extraer características texturales profundas en lugar de depender de los valores absolutos de oscuridad de los píxeles.
3. Para resolver el límite físico de los sensores ópticos frente al relieve andino extremo, se recomienda imperativamente que las futuras arquitecturas U-Net incorporen una capa de entrada matricial que no se limite a las bandas multiespectrales (RGB, NIR, SWIR). Se debe añadir un canal adicional que contenga el Modelo Digital de Elevación (DEM) o un mapa de pendientes derivado de este. Al dotar a la red neuronal de esta variable morfométrica, el algoritmo adquirirá un contexto geométrico tridimensional, permitiéndole discernir matemáticamente que las zonas de pendientes abruptas no pueden albergar cuerpos de agua estancada, solucionando de raíz el problema de los falsos positivos por sombras topográficas.
4. Para mitigar el error residual provocado por los píxeles mixtos (mixels) en los bordes irregulares de las lagunas y para validar la detección de pequeños

ecosistemas lénticos (ccochas), se aconseja ejecutar campañas de fotogrametría con Vehículos Aéreos No Tripulados (VANTs) de forma simultánea al paso del satélite Sentinel-2. Obtener ortomosaicos centimétricos en el mismo rango de fecha y hora permitirá construir máscaras muy precisas, mejorando el entrenamiento de la red y estableciendo un catastro hídrico local de altísima fidelidad.

5.3. Trabajos futuros.

1. Dado que el hallazgo limitante más significativo de la investigación fue la clasificación errónea de las densas sombras orográficas andinas como cuerpos de agua (generando 53,329 píxeles de falsos positivos en el conjunto de prueba), el principal trabajo futuro debe enfocarse en modificar la capa de entrada de la actual arquitectura U-Net. En lugar de cambiar de red neuronal, se propone incorporar un canal adicional que contenga el Modelo Digital de Elevación (DEM) o un mapa de pendientes derivado. Al fusionar las bandas espectrales ópticas con esta variable morfométrica, el algoritmo adquirirá un contexto geométrico tridimensional, permitiéndole discernir matemáticamente que las zonas con pendientes abruptas no pueden albergar masas de agua estancada, solucionando el problema de las sombras sin requerir arquitecturas más complejas.
2. A partir de las deficiencias detectadas en la fase de preprocesamiento, se sugiere que futuras investigaciones optimicen el conjunto de datos de entrenamiento sin alterar el modelo computacional. Esto implica implementar la retención de ejemplos negativos difíciles (hard negatives), es decir, conservar deliberadamente parches de imágenes que contengan quebradas oscuras y relieve escarpado sin agua, para que la red aprenda a contrastarlos. Asimismo, se debe ampliar la técnica de aumentación de datos incorporando transformaciones radiométricas

(como variaciones aleatorias de brillo, ajustes de contraste o simulación de ruido espectral). Estas perturbaciones obligarán a los filtros convolucionales a diferenciar el agua de la sombra basándose en sus texturas profundas y no únicamente en el nivel de oscuridad del píxel.

3. Tras haber validado que la arquitectura U-Net posee una excelente capacidad para delimitar los contornos irregulares con un IoU de 0.8739, el siguiente paso lógico es escalar la aplicación del modelo para resolver el problema geomorfológico de los píxeles mixtos (mixels) en las orillas. Se propone entrenar y evaluar esta misma arquitectura utilizando imágenes de ultra-alta resolución espacial, tales como ortomosaicos generados por Vehículos Aéreos No Tripulados (VANTs) o constelaciones satelitales comerciales sub-métricas. Esta transición tecnológica directa permitiría que el modelo pase de identificar lagunas a segmentar semánticamente infraestructuras hídricas de microescala, como canales de riego menores a 5 metros de ancho.

REFERENCIAS BIBLIOGRÁFICAS.

- Autoridad Nacional del Agua. (2021). Política y estrategia nacional de recursos hídricos. Ministerio de Desarrollo Agrario y Riego.
<https://www.gob.pe/institucion/ana/informes-publicaciones/>
- Centro Nacional de Estimación, Prevención y Reducción del Riesgo de Desastres. (2023). Escenario de riesgo por déficit hídrico ante el posible Fenómeno El Niño para el periodo septiembre - noviembre 2023. Ministerio de Defensa.
<https://sigrid.cenepred.gob.pe/sigridv3/documento/15748>
- Chen, Y., Fan, R., Yang, X., Wang, J., & Latif, A. (2018). Extraction of urban water bodies from high-resolution remote-sensing imagery using deep learning. *Water*, 10(5), 585. <https://doi.org/10.3390/w10050585>
- Chuvieco, E. (2020). Teledetección ambiental: La observación de la Tierra desde el espacio. Ediciones Rialp.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179–211.
https://doi.org/10.1207/s15516709cog1402_1
- Feng, W., Sui, H., Huang, W., Xu, C., & An, K. (2019). Water body extraction from very high-resolution remote sensing imagery using deep U-Net and a superpixel-based conditional random field model. *IEEE Geoscience and Remote Sensing Letters*, 16(4), 618-622. <https://doi.org/10.1109/LGRS.2018.2879492>
- Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, 140, 23–35.
<https://doi.org/10.1016/j.rse.2013.08.029>
- García Tadeo, D. A. (2023). Detección de avalanchas de origen glaciar a través de cambios del flujo de radón en el ambiente proglacial Palcacocha [Tesis de Título Profesional, Universidad Nacional Santiago Antúnez de Mayolo]. Repositorio Institucional UNASAM.

- Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. En Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (pp. 249-256).
- Gobierno Regional de Ayacucho. (2022). Plan de Desarrollo Regional Concertado Ayacucho al 2030. Gerencia Regional de Planeamiento, Presupuesto y Acondicionamiento Territorial.
<https://www.gob.pe/institucion/regionayacucho/informes-publicaciones/>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
<http://www.deeplearningbook.org>
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. En Proceedings of the IEEE International Conference on Computer Vision (ICCV), 1026-1034.
<https://doi.org/10.1109/ICCV.2015.123>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. d. P. (2014). *Metodología de la investigación* (6ta ed.). McGraw-Hill Education.
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Instituto Nacional de Investigación en Glaciares y Ecosistemas de Montaña. (2018). Inventario nacional de glaciares y lagunas de origen glaciar 2018. Ministerio del Ambiente. <https://hdl.handle.net/20.500.12816/5053>
- Intergovernmental Panel on Climate Change. (2023). Climate Change 2023: Synthesis Report. Summary for Policymakers. IPCC. <https://www.ipcc.ch>
- Isikdogan, F., Bovik, A. C., & Passalacqua, P. (2017). Surface water mapping by deep learning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(11), 4909-4918.
<https://doi.org/10.1109/JSTARS.2017.2735443>

- Jaderberg, M., Simonyan, K., Zisserman, A., & Kavukcuoglu, K. (2015). Spatial Transformer Networks. En *Advances in Neural Information Processing Systems*, 2017–2025.
- Jensen, J. R. (2016). *Introductory digital image processing: A remote sensing perspective* (4ta ed.). Pearson.
- Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. *Conferencia Internacional sobre Representaciones de Aprendizaje (ICLR)*. <https://arxiv.org/abs/1412.6980>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>
- Li, Z., & Leong, W. (2023). Super-resolution deep neural networks for water classification from free multispectral satellite imagery. *Journal of Hydrology*, 626, Artículo 130248. <https://doi.org/10.1016/j.jhydrol.2023.130248>
- Lillesand, T., Kiefer, R. W., & Chipman, J. (2015). *Remote sensing and image interpretation* (7ma ed.). John Wiley & Sons.
- Lin, M., Lucas, H. C., & Shmueli, G. (2011). Research commentary—Too big to fail: Large samples and the p-value problem. *Information Systems Research*, 22(1), 22–27. <https://doi.org/10.1287/isre.1090.0244>
- Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. En *Proceedings of the International Conference on Machine Learning (ICML)*, 30, 3.

- Mancilla Fuentes, N. (2021). Detección de Cambios en Cuerpos de Agua Mediante Imágenes Satelitales Multitemporales [Proyecto de titulación, Universidad de Valparaíso]. Repositorio Institucional de la Universidad de Valparaíso.
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425–1432. <https://doi.org/10.1080/01431169608948714>
- Ministerio del Ambiente. (2022). Plan nacional de adaptación al cambio climático del Perú: Un insumo para la actualización de la estrategia nacional ante el cambio climático. MINAM. <https://www.gob.pe/institucion/minam/informes-publicaciones/>
- Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. En *Proceedings of the 27th International Conference on Machine Learning (ICML)*.
- Palma Gongora, Y. A. H. (2023). Estudio de la evolución temporal de la concentración de coliformes fecales en aguas marinas usando imágenes satelitales en las inmediaciones de la Bahía Pucusana y la PTAR Chira, Lima – Perú [Tesis de Título Profesional, Universidad Peruana de Ciencias Aplicadas]. Repositorio Académico UPC.
- Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633), 418–422. <https://doi.org/10.1038/nature20584>
- Presidencia del Consejo de Ministros. (19 de septiembre de 2023). Decreto Supremo que declara el Estado de Emergencia en varios distritos... por peligro inminente

ante déficit hídrico. (Decreto Supremo N° 104-2023-PCM). Diario Oficial El Peruano.

Ramos Barzola, M. A. (2021). Estimación de la pérdida de cobertura glaciar al Noreste de la Cordillera Blanca– Perú, haciendo uso del producto Landsat TM, ETM+, OLI-TIR y Sentinel-2 MSI (1984-2021) [Tesis de Título Profesional, Universidad de Medellín]. Repositorio Institucional UdeM.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. En N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234–241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>

United Nations World Water Assessment Programme. (2023). The United Nations World Water Development Report 2023: Partnerships and Cooperation for Water. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000384655>

Xu, H. (2006). Modification of normalised difference water index (MNDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025–3033. <https://doi.org/10.1080/01431160600589179>

Zhang, X., & Li, J. (2022). MRSE-Net: Multiscale Residuals and SE-Attention Network for Water Body Segmentation From Satellite Images. *IEEE Journal of*

Selected Topics in Applied Earth Observations and Remote Sensing, 15, 5049–5064. <https://doi.org/10.1109/JSTARS.2022.3185245>

Zhang, J., & Morris, A. J. (1998). A Sequential Learning Approach for Single Hidden Layer Neural Networks. *Neural Networks*, 11(1), 65–80.
[https://doi.org/10.1016/S0893-6080\(97\)00111-1](https://doi.org/10.1016/S0893-6080(97)00111-1)

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4), 8–36.
<https://doi.org/10.1109/MGRS.2017.2762307>

ANEXO A

Algoritmos

Algorithm 1: Entrenamiento del modelo Deep Learning (U-Net)

Entrada: Imágenes X y Máscaras Y (formato .tif), $\alpha = 1e^{-4}$, $N = 100$
Salida: Modelo entrenado M_{opt}

```
// --- FASE 1: PREPROCESAMIENTO ---
1 for cada imagen  $x_i$  en  $X$  do
2    $x_i \leftarrow \text{LeerTiff}(x_i)$ 
   // Normalización Min-Max (Robusta)
3    $x_i \leftarrow \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i) + \epsilon}$ 
4    $y_i \leftarrow (y_i > 0) ? 1 : 0$ 
5 end
// --- FASE 2: CONFIGURACIÓN ---
6  $Train, Val \leftarrow \text{Split}(X, Y, 0.85)$ 
7  $Gen \leftarrow \text{DataAugmentation}(Train, \text{Rot}=20, \text{Zoom}=0.1, \text{Flip}=\text{True})$ 
   // U-Net Modificada: 32 filtros, LeakyReLU, L2
8  $M \leftarrow \text{BuildUNet}(\text{Filters}=32, \text{Activ}=\text{'LeakyReLU'}, \text{Reg}=\text{'L2'})$ 
9  $Optim \leftarrow \text{Adam}(\text{lr}=\alpha)$ 
// --- FASE 3: BUCLE DE ENTRENAMIENTO ---
10 while  $epoch < N$  do
11    $Batch \leftarrow \text{next}(Gen)$ 
12    $Pred \leftarrow M(Batch)$ 
   // Loss Híbrida: BCE Ponderada + Dice Loss
13    $L \leftarrow \text{WeightedBCE}(Y, Pred) + \text{DiceLoss}(Y, Pred)$ 
14   Actualizar pesos  $\theta$  usando  $\nabla L$ 
15   if  $Val\_Loss$  mejora then
16     GuardarModelo( $M$ )
17      $paciencia \leftarrow 0$ 
18   end
19   else
20      $paciencia \leftarrow paciencia + 1$ 
21     if  $paciencia > 15$  then
22       break (Early Stopping)
23     end
24   end
25 end
26 return  $M_{opt}$ 
```

Algorithm 2: Validación del Modelo y Pruebas Estadísticas

Entrada: Dataset Prueba X_{test}, Y_{test} , Modelo Entrenado M , Nivel $\alpha = 0,05$

Salida: Métricas de Desempeño y Resultado Estadístico

```
// 1. Cargar Modelo con Objetos Personalizados
1  $Objetos \leftarrow \{ 'iou\_metric', 'dice\_loss', 'total\_loss' \}$ 
2  $M \leftarrow LoadModel('modelo.keras', custom\_objects = Objetos)$ 
3  $Y_{pred\_total} \leftarrow [], Y_{real\_total} \leftarrow []$ 
// 2. Bucle de Inferencia
4 for cada imagen  $x_i$  en  $X_{test}$  do
5    $ProbMap \leftarrow M.predict(x_i)$ 
6    $BinMap \leftarrow (ProbMap > 0,5)?1:0$ 
7    $Y_{pred\_total}.append(Flatten(BinMap))$ 
8    $Y_{real\_total}.append(Flatten(Y_{test}[i]))$ 
9 end
// 3. Cálculo de Matriz de Confusión Global
10  $MC \leftarrow ConfusionMatrix(Y_{real\_total}, Y_{pred\_total})$ 
// 4. Prueba de McNemar
// Calcula si la discordancia entre errores es significativa
11  $b \leftarrow MC[0,1]$  // Falsos Positivos
12  $c \leftarrow MC[1,0]$  // Falsos Negativos
13  $Estadistico \leftarrow \frac{(b-c)^2}{b+c}$ 
14  $Valor\_P \leftarrow Chi2.CDF(Estadistico, 1)$ 
15 if  $Valor\_P < \alpha$  then
16   return "Diferencia Significativa ( $H1$ )"
17 end
18 else
19   return "Sin Diferencia Significativa ( $H0$ )"
20 end
```

ANEXO B

Diagramas de flujo

Diagrama de flujo del Entrenamiento del modelo

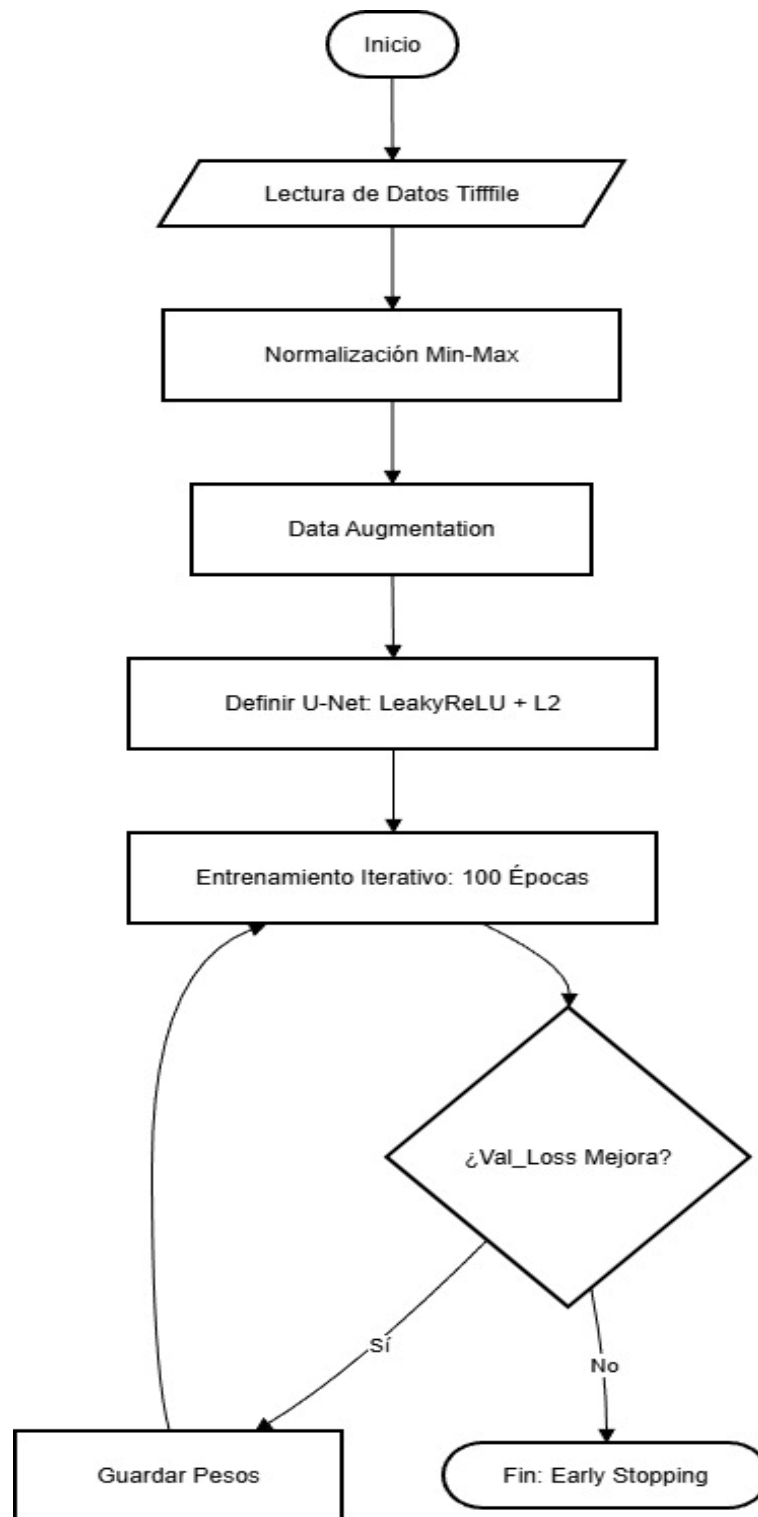
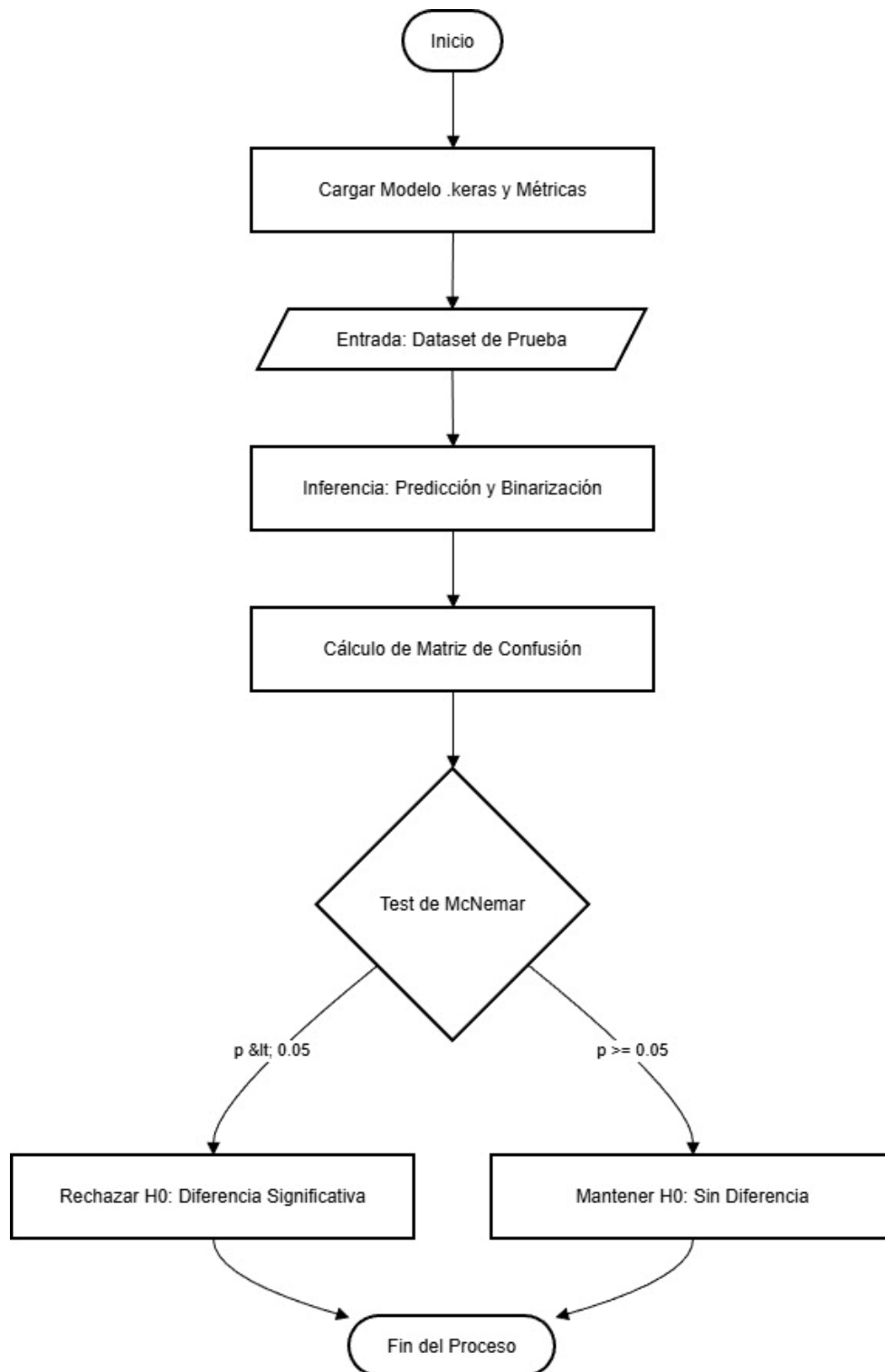


Diagrama de flujo del prueba del modelo y pruebas estadísticas



ANEXO C

Script del modelo

```

1 import os
2 import cv2
3 import numpy as np
4 import tiff as tiff
5 import tensorflow as tf
6 from tensorflow.keras import layers, models, regularizers
7 from tensorflow.keras import backend as K
8 from sklearn.model_selection import train_test_split
9 from tensorflow.keras.preprocessing.image import ImageDataGenerator
10 import math
11
12 # --- 1. CONFIGURACIÓN ---
13 ANCHO_IMG = 256
14 ALTO_IMG = 256
15 CHANNELS = 3
16 BATCH_SIZE = 32
17 EPOCHS = 100
18 RUTA_BASE = '/content/drive/MyDrive/_TESIS_UNSCH_2025/DATASETS/'
19
20 # --- 2. MÉTRICAS Y FUNCIONES DE PÉRDIDA PERSONALIZADAS ---
21 def iou_metric(y_true, y_pred):
22     y_true = K.flatten(y_true)
23     y_pred = K.flatten(K.cast(y_pred > 0.5, 'float32'))
24     intersection = K.sum(y_true * y_pred)
25     union = K.sum(y_true) + K.sum(y_pred) - intersection
26     return (intersection + 1e-7) / (union + 1e-7)
27
28 def dice_loss(y_true, y_pred):
29     y_true = K.flatten(y_true)
30     y_pred = K.flatten(y_pred)
31     intersection = K.sum(y_true * y_pred)
32     return 1 - (2. * intersection + 1e-7) / (K.sum(y_true) + K.sum(y_pred) + 1e-7)
33
34 def total_loss(y_true, y_pred):
35     # Loss Híbrida: Binary CrossEntropy Ponderada + Dice Loss
36     epsilon = K.epsilon()
37     y_pred = tf.clip_by_value(y_pred, epsilon, 1. - epsilon)
38
39     # Nota: El peso 'pos_weight_value' se calcula dinámicamente en la carga de datos
40     # Aquí se asume el valor calculado en el script principal
41     pos_weight_value = 1.0
42     weighted_bce = - (pos_weight_value * y_true * tf.math.log(y_pred) +
43                      (1. - y_true) * tf.math.log(1. - y_pred))
44
45     return tf.reduce_mean(weighted_bce) + dice_loss(y_true, y_pred)
46
47 # --- 3. ARQUITECTURA U-NET MODIFICADA (LEAKY RELU + L2) ---
48 def construccion_unet(img_height=256, img_width=256, img_channels=3,
49                        dropout_rate=0.28, kernel_reg=0.0005):
50
51     inputs = layers.Input((img_height, img_width, img_channels))
52     l2_reg = regularizers.l2(kernel_reg)
53     activation = layers.LeakyReLU(negative_slope=0.1)
54
55     def conv_block(input_tensor, num_filters):
56         x = layers.Conv2D(num_filters, (3, 3), padding="same", kernel_regularizer=l2_reg)(input_tensor)
57         x = layers.BatchNormalization()(x)
58         x = activation(x)
59         if dropout_rate > 0:
60             x = layers.Dropout(dropout_rate)(x)
61         x = layers.Conv2D(num_filters, (3, 3), padding="same", kernel_regularizer=l2_reg)(x)
62         x = layers.BatchNormalization()(x)
63         x = activation(x)
64         return x
65
66     # Encoder (Bajada)
67     c1 = conv_block(inputs, 32)
68     p1 = layers.MaxPooling2D((2, 2))(c1)
69
70     c2 = conv_block(p1, 64)
71     p2 = layers.MaxPooling2D((2, 2))(c2)
72
73     c3 = conv_block(p2, 128)
74     p3 = layers.MaxPooling2D((2, 2))(c3)
75
76     c4 = conv_block(p3, 256)
77     p4 = layers.MaxPooling2D((2, 2))(c4)

```

```

78
79 # Bridge (Cuello de botella)
80 c5 = conv_block(p4, 512)
81
82 # Decoder (Subida)
83 u6 = layers.Conv2DTranspose(256, (2, 2), strides=(2, 2), padding="same")(c5)
84 u6 = layers.concatenate([u6, c4])
85 c6 = conv_block(u6, 256)
86
87 u7 = layers.Conv2DTranspose(128, (2, 2), strides=(2, 2), padding="same")(c6)
88 u7 = layers.concatenate([u7, c3])
89 c7 = conv_block(u7, 128)
90
91 u8 = layers.Conv2DTranspose(64, (2, 2), strides=(2, 2), padding="same")(c7)
92 u8 = layers.concatenate([u8, c2])
93 c8 = conv_block(u8, 64)
94
95 u9 = layers.Conv2DTranspose(32, (2, 2), strides=(2, 2), padding="same")(c8)
96 u9 = layers.concatenate([u9, c1])
97 c9 = conv_block(u9, 32)
98
99 outputs = layers.Conv2D(1, (1, 1), activation='sigmoid')(c9)
100 return models.Model(inputs=[inputs], outputs=[outputs])
101
102 # --- 4. CARGA DE DATOS ---
103 def cargar_datos(ruta):
104     imagenes, mascarar = [], []
105     archivos_img = sorted([f for f in os.listdir(ruta + 'images/') if f.endswith('.tif')])
106
107     for nombre in archivos_img:
108         try:
109             ruta_img = os.path.join(ruta, 'images/', nombre)
110             ruta_mask = os.path.join(ruta, 'labels/', nombre)
111             if not os.path.exists(ruta_mask): continue
112
113             # Lectura TIFF y Resize
114             img = tiff.imread(ruta_img)
115             mask = tiff.imread(ruta_mask)
116
117             if img.shape[:2] != (ALTO_IMG, ANCHO_IMG):
118                 img = cv2.resize(img, (ANCHO_IMG, ALTO_IMG))
119                 mask = cv2.resize(mask, (ANCHO_IMG, ALTO_IMG), interpolation=cv2.INTER_NEAREST)
120
121             # Normalización Robusta Min-Max por imagen
122             img = img.astype(np.float32)
123             min_val, max_val = np.min(img), np.max(img)
124             if max_val - min_val > 1e-7:
125                 img = (img - min_val) / (max_val - min_val)
126             else:
127                 img = np.zeros_like(img)
128
129             mask = (mask > 0).astype(np.float32) # Binarización
130             imagenes.append(img)
131             mascarar.append(mask)
132         except Exception as e:
133             continue
134
135     return np.array(imagenes), np.expand_dims(np.array(mascarar), axis=3)
136
137 # --- 5. EJECUCIÓN PRINCIPAL ---
138 if __name__ == "__main__":
139     # A. Cargar y Dividir
140     X, Y = cargar_datos(RUTA_BASE)
141     X_train, X_val, Y_train, Y_val = train_test_split(X, Y, test_size=0.15, random_state=42)
142
143     # B. Generadores (Data Augmentation)
144     data_gen_args = dict(rotation_range=20, width_shift_range=0.1, height_shift_range=0.1,
145                          zoom_range=0.1, horizontal_flip=True, vertical_flip=True)
146
147     train_gen = zip(ImageDataGenerator(**data_gen_args).flow(X_train, batch_size=BATCH_SIZE, seed=42),
148                    ImageDataGenerator(**data_gen_args).flow(Y_train, batch_size=BATCH_SIZE, seed=42))
149
150     val_gen = zip(ImageDataGenerator().flow(X_val, batch_size=BATCH_SIZE, seed=42),
151                  ImageDataGenerator().flow(Y_val, batch_size=BATCH_SIZE, seed=42))
152
153     # C. Construir Modelo
154     model = construccion_unet(dropout_rate=0.28, kernel_reg=0.0005)
155     model.compile(optimizer='adam', loss=total_loss, metrics=[iou_metric, 'accuracy'])

```

```
156
157 # D. Entrenar
158 history = model.fit(
159     train_gen,
160     steps_per_epoch=len(X_train) // BATCH_SIZE,
161     validation_data=val_gen,
162     validation_steps=math.ceil(len(X_val) / BATCH_SIZE),
163     epochs=EPOCHS
164 )
165
166 # E. Guardar
167 model.save('modelo_entrenado_unet_modificado.keras')
```

```

1 import os
2 import cv2
3 import numpy as np
4 import tensorflow as tf
5 from tensorflow.keras.models import load_model
6 from tensorflow.keras import backend as K
7 from sklearn.metrics import confusion_matrix
8 # Librería para prueba estadística
9 from statsmodels.stats.contingency_tables import mcnemar
10
11 """
12 =====
13 ANEXO 4: SCRIPT DE PRUEBA Y VALIDACIÓN ESTADÍSTICA
14 Descripción: Carga del modelo entrenado, inferencia en datos de prueba,
15             cálculo de métricas (IoU) y Test de McNemar vs MNDWI.
16 =====
17 """
18
19 # --- 1. DEFINICIÓN DE MÉTRICAS PERSONALIZADAS (Necesario para cargar el modelo) ---
20 def iou_metric(y_true, y_pred):
21     y_true = K.flatten(y_true)
22     y_pred = K.flatten(K.cast(y_pred > 0.5, 'float32'))
23     intersection = K.sum(y_true * y_pred)
24     union = K.sum(y_true) + K.sum(y_pred) - intersection
25     return (intersection + 1e-7) / (union + 1e-7)
26
27 def dice_loss(y_true, y_pred):
28     y_true = K.flatten(y_true)
29     y_pred = K.flatten(y_pred)
30     intersection = K.sum(y_true * y_pred)
31     return 1 - (2. * intersection + 1e-7) / (K.sum(y_true) + K.sum(y_pred) + 1e-7)
32
33 def total_loss(y_true, y_pred):
34     # Loss Híbrida usada en el entrenamiento
35     pos_weight = 1.0
36     epsilon = K.epsilon()
37     y_pred = tf.clip_by_value(y_pred, epsilon, 1. - epsilon)
38     bce = - (pos_weight * y_true * tf.math.log(y_pred) + (1. - y_true) * tf.math.log(1. - y_pred))
39     return tf.reduce_mean(bce) + dice_loss(y_true, y_pred)
40
41 # --- 2. CARGA DEL MODELO Y DATOS ---
42 RUTA_MODELO = 'modelo_entrenado_unet_modificado.keras'
43 RUTA_TEST_IMGS = 'dataset/test/images/'
44 RUTA_TEST_MASKS = 'dataset/test/labels/'
45
46 print(f"Cargando modelo desde: {RUTA_MODELO}")
47 # Se debe pasar el diccionario 'custom_objects' porque usamos métricas propias
48 modelo = load_model(RUTA_MODELO, custom_objects={
49     'iou_metric': iou_metric,
50     'dice_loss': dice_loss,
51     'total_loss': total_loss
52 })
53
54 # --- 3. INFERENCIA Y EVALUACIÓN ---
55 def evaluar_modelo(modelo, ruta_imgs, ruta_masks):
56     archivos = sorted(os.listdir(ruta_imgs))
57     y_real_total = []
58     y_pred_total = []
59
60     print("Iniciando inferencia en set de prueba...")
61     for nombre in archivos:
62         try:
63             # Lectura y Preprocesamiento
64             img = cv2.imread(os.path.join(ruta_imgs, nombre))
65             img = cv2.resize(img, (256, 256))
66             img = img.astype(np.float32) / 255.0 # Normalización básica para test
67             img_input = np.expand_dims(img, axis=0)
68
69             # Ground Truth
70             mask = cv2.imread(os.path.join(ruta_masks, nombre), cv2.IMREAD_GRAYSCALE)
71             mask = cv2.resize(mask, (256, 256))
72             mask = (mask > 0).astype(np.uint8) # 0 o 1
73
74             # Predicción
75             pred = modelo.predict(img_input, verbose=0)[0]
76             pred_bin = (pred > 0.5).astype(np.uint8).squeeze()
77

```

```

78         y_real_total.extend(mask.flatten())
79         y_pred_total.extend(pred_bin.flatten())
80
81     except Exception as e:
82         continue
83
84     return np.array(y_real_total), np.array(y_pred_total)
85
86 # Ejecutar evaluación
87 y_true, y_pred = evaluar_modelo(modelo, RUTA_TEST_IMGS, RUTA_TEST_MASKS)
88
89 # --- 4. PRUEBA ESTADÍSTICA DE McNEMAR (Modelo vs Referencia) ---
90 print("\n--- Ejecutando Test de McNemar ---")
91 # Tabla de contingencia: [ [Aciertos_Ambos, FalloModelo_AciertoRef],
92 #                           [AciertoModelo_FalloRef, Fallos_Ambos] ]
93 # Nota: Para McNemar simplificado, comparamos aciertos/errores binarios
94 tabla_mcnemar = confusion_matrix(y_true, y_pred)
95
96 resultado = mcnemar(tabla_mcnemar, exact=True)
97
98 print(f"Estadístico: {resultado.statistic}")
99 print(f"Valor p: {resultado.pvalue}")
100
101 alpha = 0.05
102 if resultado.pvalue < alpha:
103     print("Resultado: Se rechaza H0. Existe diferencia significativa entre el Modelo y la referencia.")
104 else:
105     print("Resultado: No se rechaza H0. No hay diferencia significativa.")

```


ANEXO D
Panel Fotográfico

Figura 01: Cuerpo de agua identificada como Laguna 01



Figura 02: Vista panormica de cuerpo de agua identificada como Laguna 01

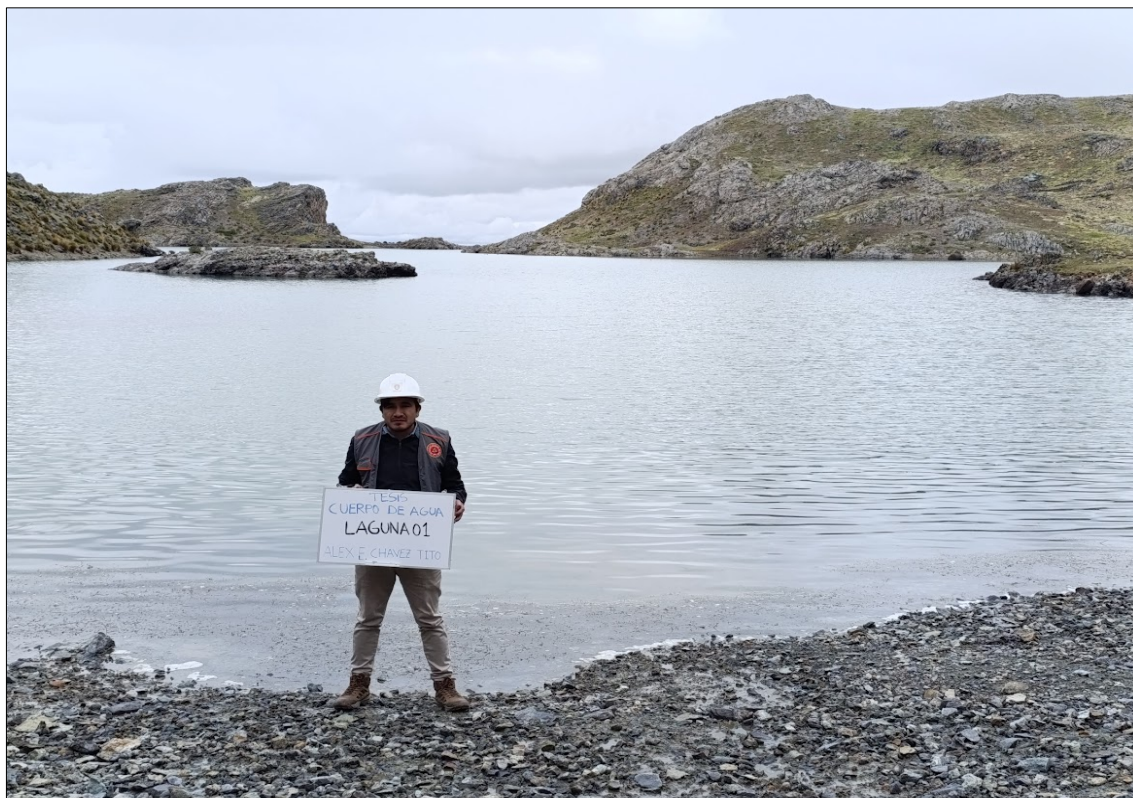


Figura 03: Pequeños cuerpos de agua avistadas en la zona de estudio



Figura 04: Vista panoramica de la laguna Yanacocha



Figura 05: Registro de la laguna Yanacocha como cuerpo de agua



Figura 06: Registro de la laguna Huatay como cuerpo de agua



Figura 07: Vista panramica de la laguna Huatay



Figura 08: Cuerpo de agua cercana a la via registrada como Laguna 04.



Figura 09: Cuerpo de agua registrada como Laguna 05 cerca a la via.

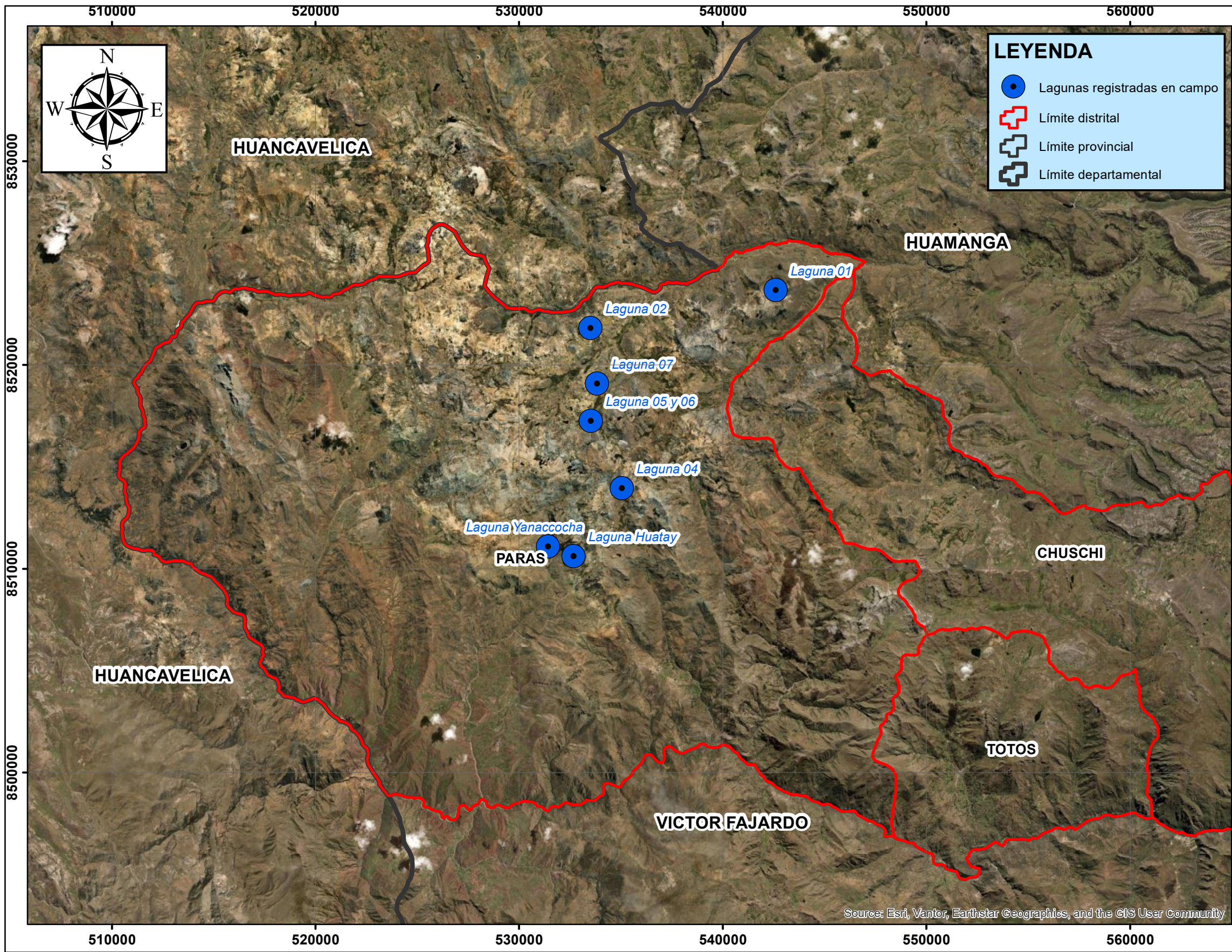


Figura 10: Cuerpo de agua cercana a la via registrada como Laguna 06.



ANEXO E

Cartografía de Lagunas Registradas en Campo



ESCALA GRÁFICA

ESCALA NUMÉRICA

1:200,000

ESPECIFICACIONES GEODÉSICAS

ELIPSOIDE:

SISTEMA GEOGRÁFICO MUNDIAL, 1984

SISTEMA DE PROYECCIÓN: UTM

ZONA: 18 SUR CUADRICULA: L

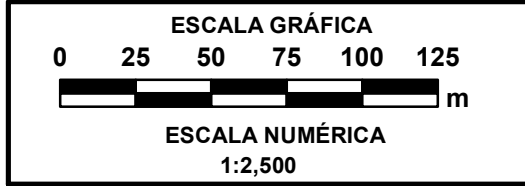
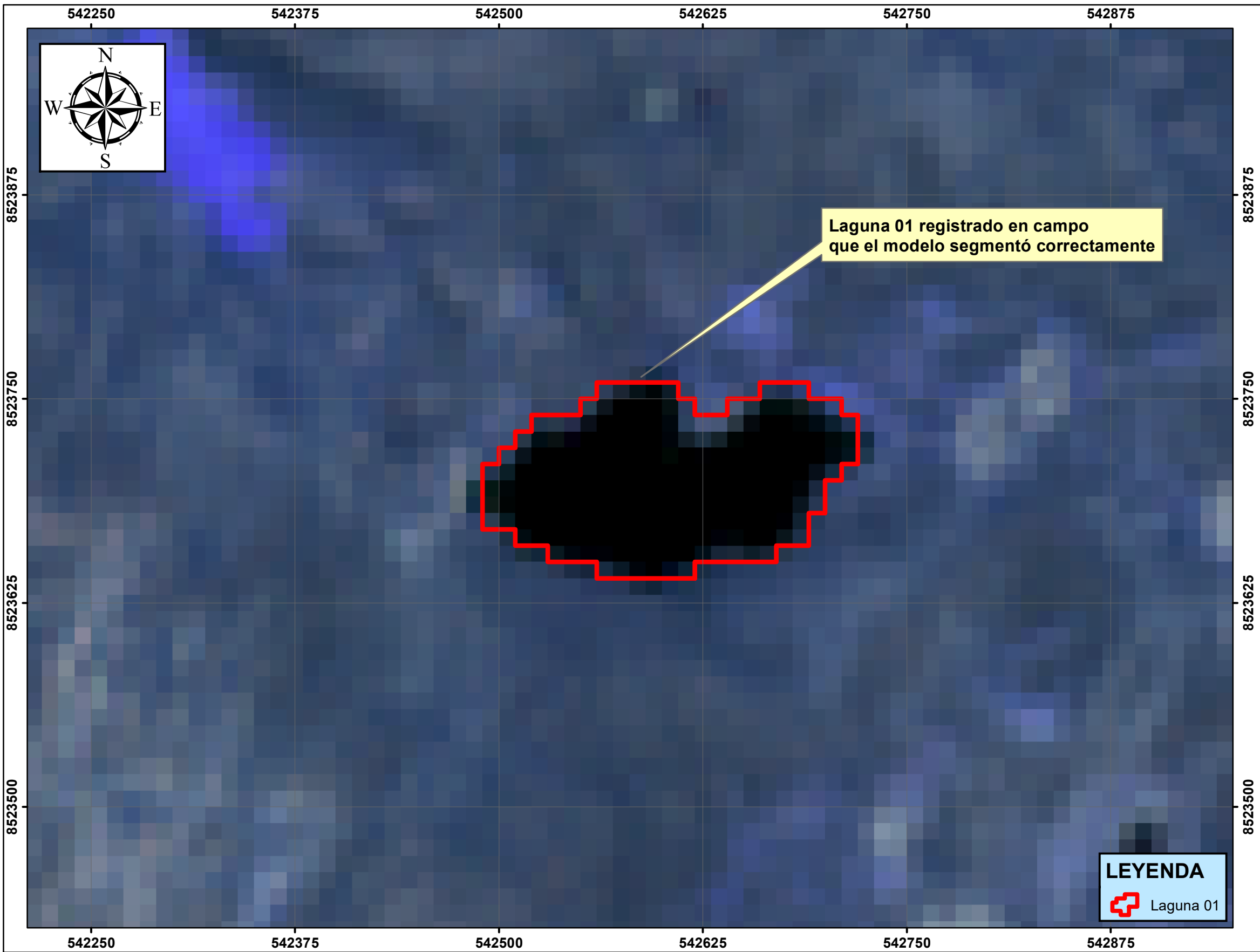
DATUM HORIZONTAL: WGS 84



TESIS:						
DETECCIÓN DE CUERPOS DE AGUA MEDIANTE EL PROCESAMIENTO DE IMÁGENES SATELITALES UTILIZANDO UN MODELO DE DEEP LEARNING EN LA PROVINCIA DE CANGALLO, 2024						
PLANO:						
UBICACION Y LOCALIZACION DE LAGUNAS REGISTRADAS EN CAMPO						
UNIVERSIDAD NACIONAL AN CRISTOBAL DE HUAMANGA	REGIÓN:	Ayacucho	DISTRITO:	Paras	ESCALA:	Indicada
	PROVINCIA:	Cangallo	LOCALIDAD:	Nevado Ritipata	FECHA:	Diciembre 2024

CÓDIGO:

UL - 01



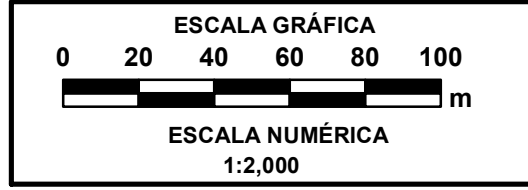
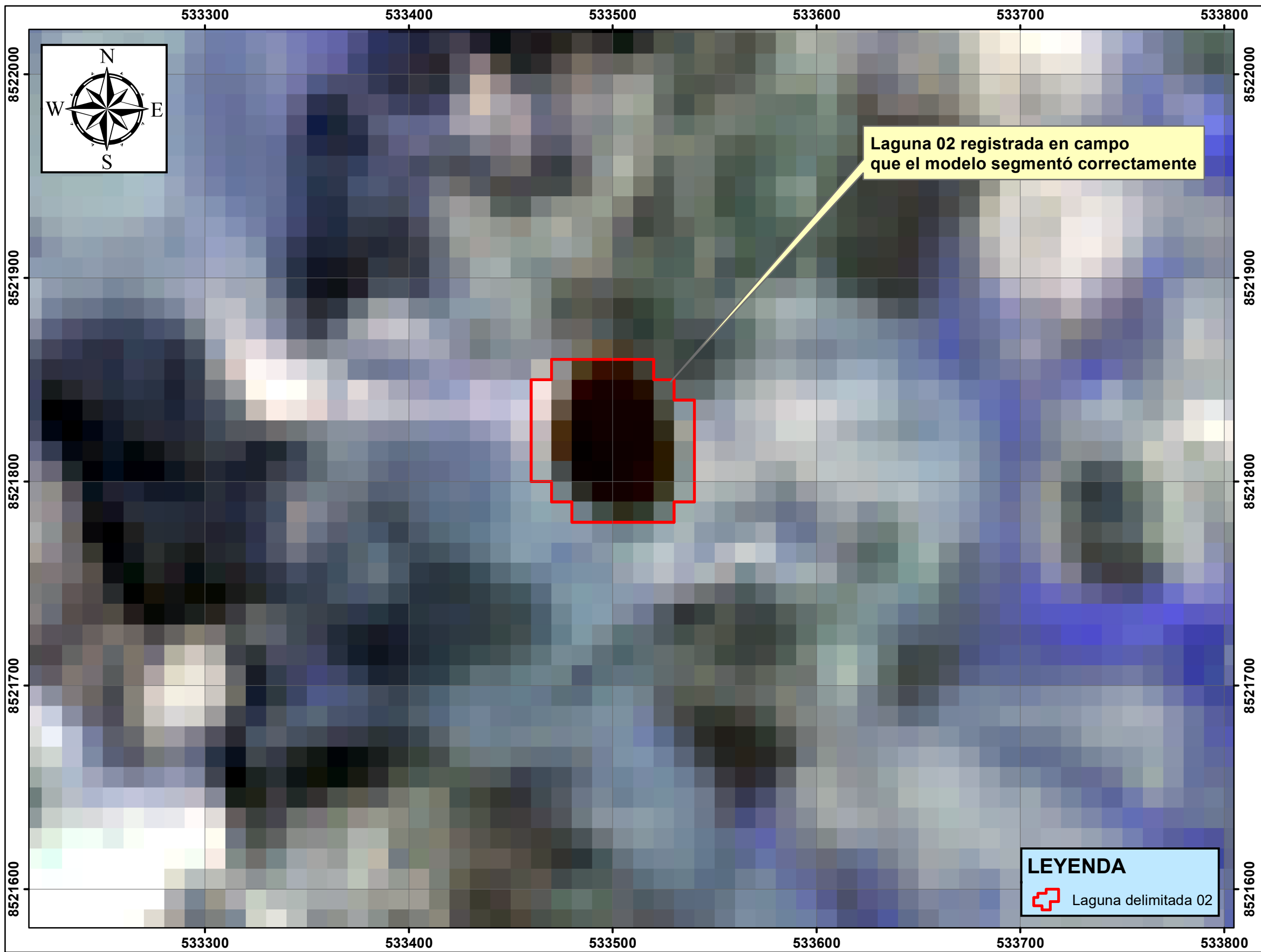
ESPECIFICACIONES GEODÉSICAS
ELIPSOIDE:
SISTEMA GEOGRAFICO MUNDIAL, 1984
SISTEMA DE PROYECCIÓN: UTM
ZONA: 18 SUR CUADRICULA: L
DATUM HORIZONTAL: WGS 84



TESIS:							
DETECCIÓN DE CUERPOS DE AGUA MEDIANTE EL PROCESAMIENTO DE IMÁGENES SATELITALES UTILIZANDO UN MODELO DE DEEP LEARNING EN LA PROVINCIA DE CANGALLO, 2024							
PLANO:							
CARTOGRAFÍA DE CUERPO DE AGUA REGISTRADO COMO LAGUNA 01							
UNIVERSIDAD NACIONAL SAN CRISTOBAL DE HUAMANGA		REGIÓN:	Ayacucho	DISTRITO:	Paras	ESCALA:	Indicada
		PROVINCIA:	Cangallo	LOCALIDAD:	Nevado Ritipata	FECHA:	Diciembre 2024

CÓDIGO:

LA - 01



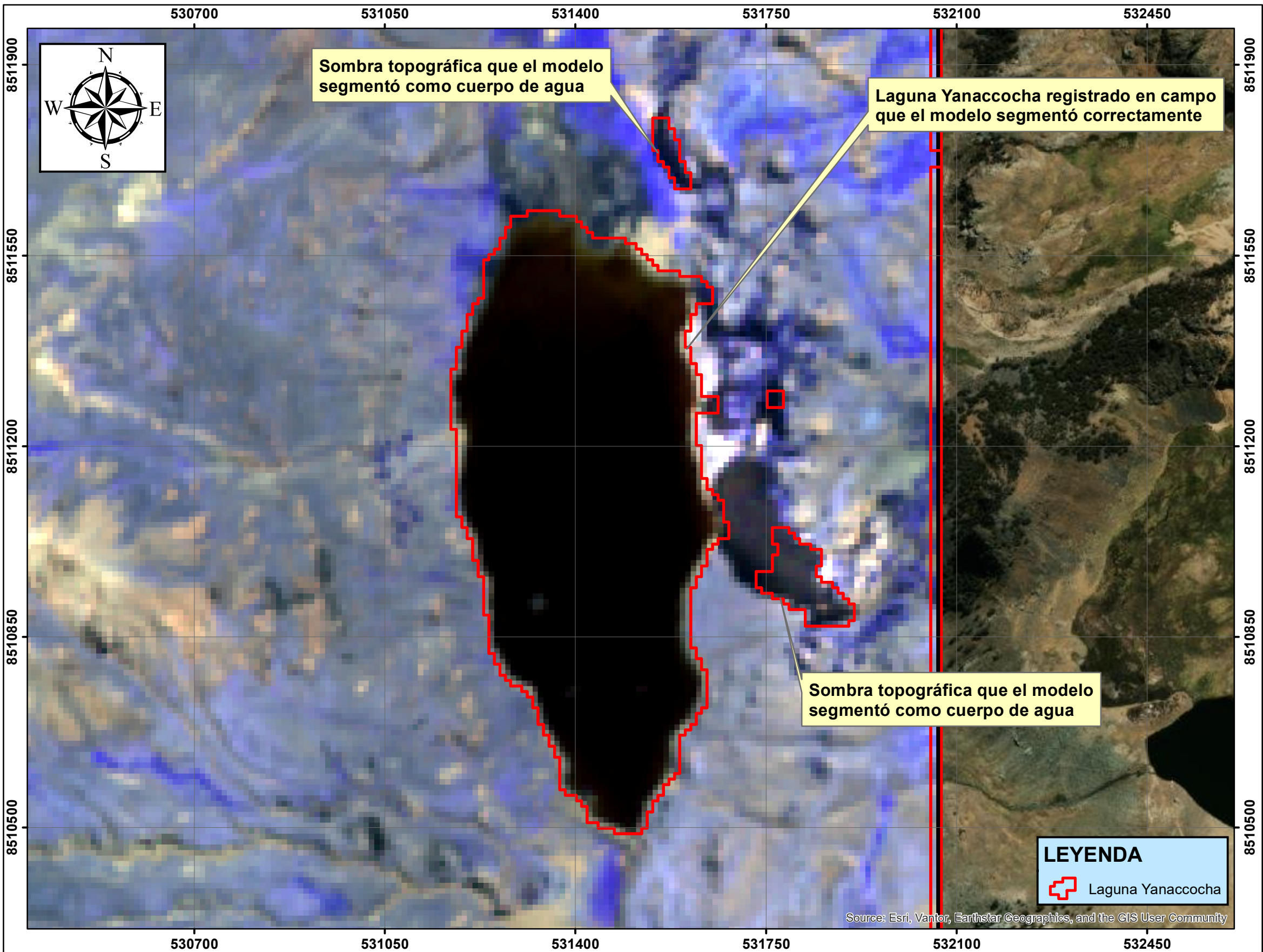
ESPECIFICACIONES GEODÉSICAS
ELIPSOIDE:
SISTEMA GEOGRAFICO MUNDIAL, 1984
SISTEMA DE PROYECCIÓN: UTM
ZONA: 18 SUR CUADRICULA: L
DATUM HORIZONTAL: WGS 84



TESIS:	DETECCIÓN DE CUERPOS DE AGUA MEDIANTE EL PROCESAMIENTO DE IMÁGENES SATELITALES UTILIZANDO UN MODELO DE DEEP LEARNING EN LA PROVINCIA DE CANGALLO, 2024					
	PLANO: CARTOGRAFÍA DE CUERPO DE AGUA REGISTRADO COMO LAGUNA 02					
UNIVERSIDAD NACIONAL SAN CRISTOBAL DE HUAMANGA	REGIÓN:	Ayacucho	DISTRITO:	Paras	ESCALA:	Indicada
	PROVINCIA:	Cangallo	LOCALIDAD:	Nevado Ritipata	FECHA:	Diciembre 2024

CÓDIGO:

LA - 02

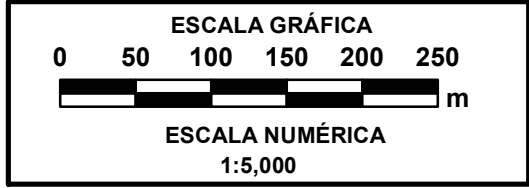
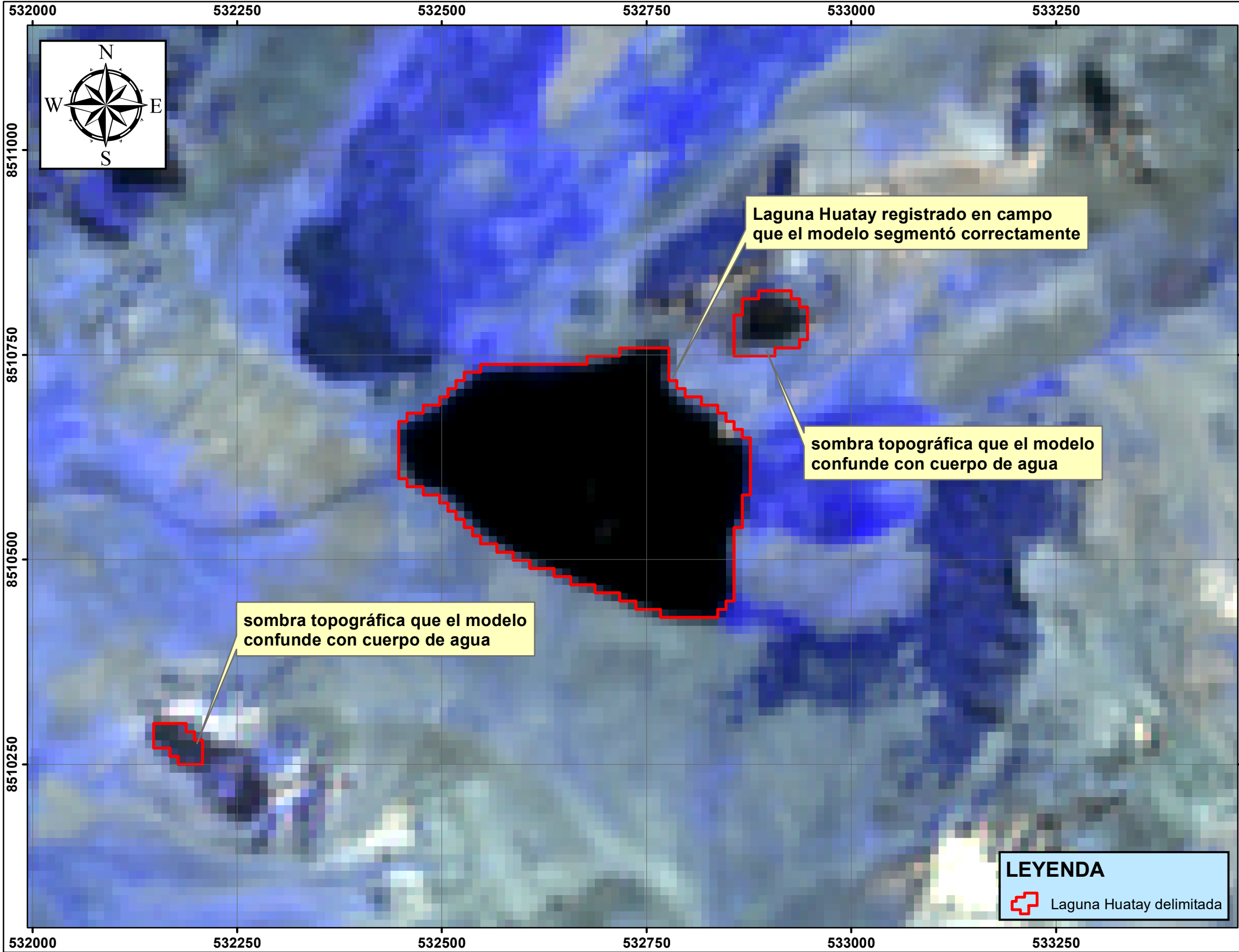


ESPECIFICACIONES GEODÉSICAS
ELIPSOIDE:
SISTEMA GEOGRÁFICO MUNDIAL, 1984
SISTEMA DE PROYECCIÓN: UTM
ZONA: 18 SUR CUADRICULA: L
DATUM HORIZONTAL: WGS 84



TESIS:				
DETECCIÓN DE CUERPOS DE AGUA MEDIANTE EL PROCESAMIENTO DE IMÁGENES SATELITALES UTILIZANDO UN MODELO DE DEEP LEARNING EN LA PROVINCIA DE CANGALLO, 2024				
PLANO:				
CARTOGRAFÍA DE CUERPO DE AGUA REGISTRADO COMO LAGUNA YANACCOCHA				
UNIVERSIDAD NACIONAL SAN CRISTOBAL DE HUAMANGA	REGIÓN:	Ayacucho	DISTRITO:	Paras
	PROVINCIA:	Cangallo	LOCALIDAD:	Nevado Ritipata
			ESCALA:	Indicada
			FECHA:	Diciembre 2024

CÓDIGO:
LA - 03



ESPECIFICACIONES GEODÉSICAS
ELIPSOIDE:
SISTEMA GEOGRAFICO MUNDIAL, 1984
SISTEMA DE PROYECCIÓN: UTM
ZONA: 18 SUR CUADRICULA: L
DATUM HORIZONTAL: WGS 84

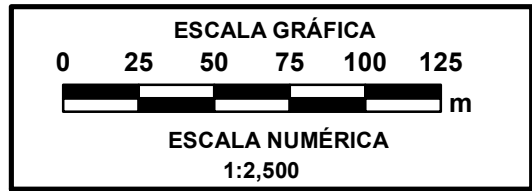
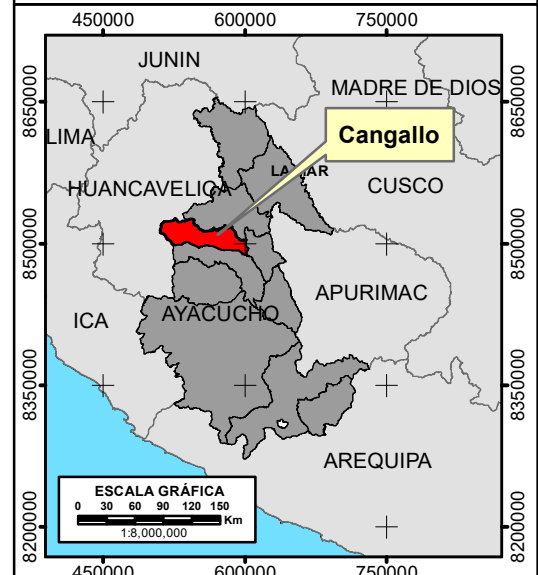
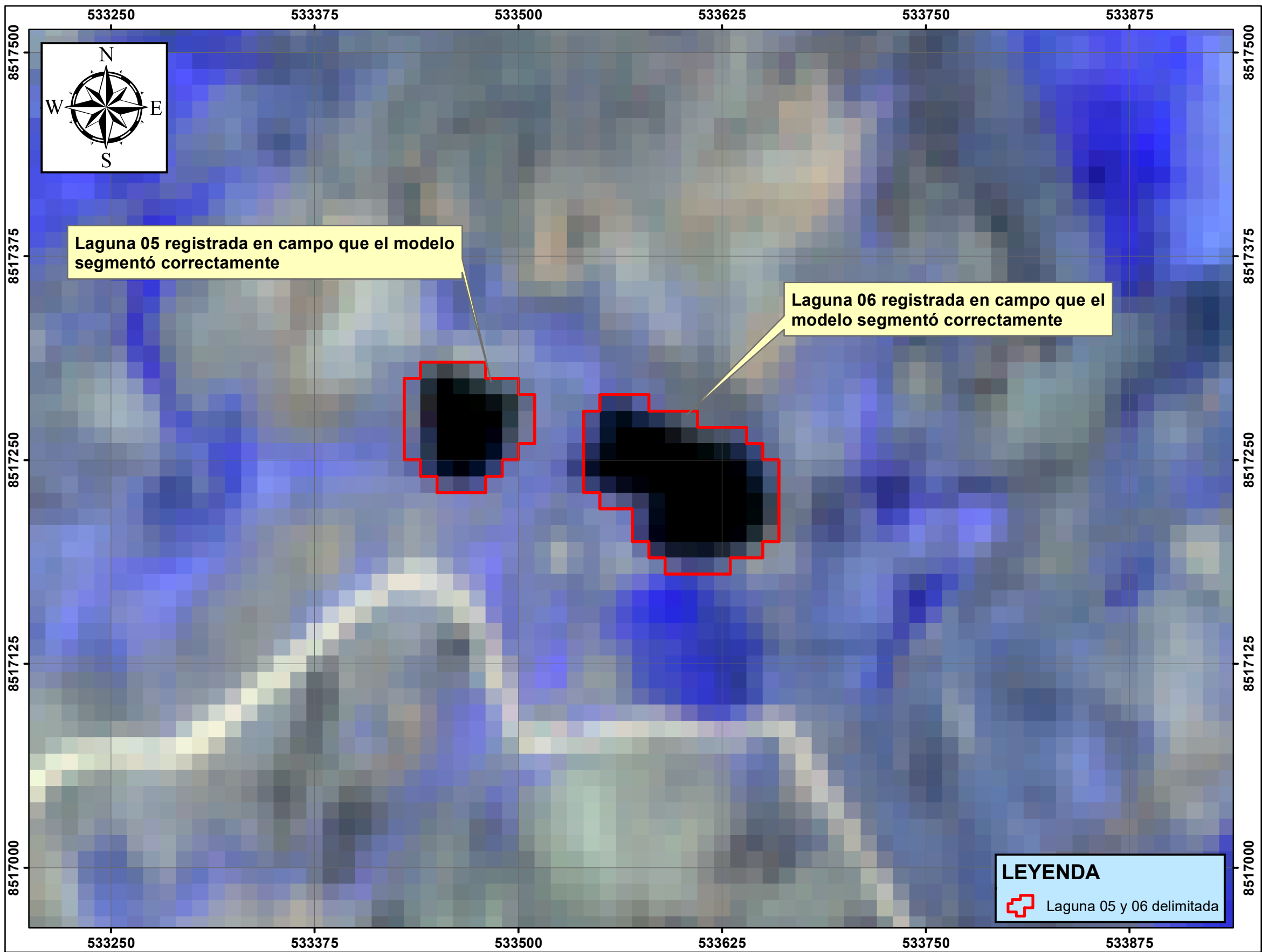


TESIS:
DETECCIÓN DE CUERPOS DE AGUA MEDIANTE EL PROCESAMIENTO DE IMÁGENES SATELITALES
UTILIZANDO UN MODELO DE DEEP LEARNING EN LA PROVINCIA DE CANGALLO, 2024

PLANO:
CARTOGRAFÍA DE CUERPO DE AGUA REGISTRADO COMO LAGUNA HUATAY

ENTIDAD:	UNIVERSIDAD NACIONAL SAN CRISTOBAL DE HUAMANGA	REGIÓN:	Ayacucho	DISTRITO:	Paras	ESCALA:	Indicada
		PROVINCIA:	Cangallo	LOCALIDAD:	Nevado Ritipata	FECHA:	Diciembre 2024

CÓDIGO:
LA - 04



ESPECIFICACIONES GEODÉSICAS
ELIPSOIDE:
SISTEMA GEOGRÁFICO MUNDIAL, 1984
SISTEMA DE PROYECCIÓN: UTM
ZONA: 18 SUR CUADRICULA: L
DATUM HORIZONTAL: WGS 84



TESIS:						
DETECCIÓN DE CUERPOS DE AGUA MEDIANTE EL PROCESAMIENTO DE IMÁGENES SATELITALES UTILIZANDO UN MODELO DE DEEP LEARNING EN LA PROVINCIA DE CANGALLO, 2024						
PLANO:						
CARTOGRAFÍA CUERPOS DE AGUA REGISTRADAS COMO LAGUNAS 05 Y 06						
UNIVERSIDAD NACIONAL SAN CRISTOBAL DE HUAMANGA	REGIÓN:	Ayacucho	DISTRITO:	Paras	ESCALA:	Indicada
	PROVINCIA:	Cangallo	LOCALIDAD:	Nevado Ritipata	FECHA:	Diciembre 2024

CÓDIGO:
LA - 06

ANEXO F

Matriz de Operacionalidad

Detección de cuerpos de agua mediante el procesamiento de imágenes satelitales utilizando un modelo de Deep Learning en la provincia de Cangallo, 2024.

PROBLEMA	OBJETIVO	HIPÓTESIS	METODOLOGÍA	VARIABLES E INDICADORES	DIMENSION	INDICADORES	TECNICAS / INSTRUMENTOS	ESCALA	UNIDAD DE MEDIDA
¿Cómo influye la implementación de un modelo de Deep Learning en la precisión de la detección automatizada de cuerpos de agua mediante el procesamiento de imágenes satelitales en la provincia de Cangallo, 2024?	Evaluar la influencia de la implementación de un modelo de Deep Learning en la precisión de la detección automatizada de cuerpos de agua mediante el procesamiento de imágenes satelitales en la provincia de Cangallo, 2024.	La implementación de un modelo de Deep Learning mejora significativamente la precisión de la detección automatizada de cuerpos de agua en comparación con métodos espectrales tradicionales en la provincia de Cangallo.	Tipo de Investigación: investigación aplicada (genera beneficio y bienestar a la sociedad). Enfoque: Cuantitativo (Ya que esta se basa en la toma de los datos se medirá la variable mediante instrumentos). Alcance (o nivel): Explicativo (no se limita a observar la ubicación del agua, si no que busca explicar como la manipulación de la arquitectura y el preprocesamiento determinan la precisión final). Diseño no experimental: No se manipula el fenómeno real, sino que se analiza imágenes satelitales existentes. Incluye un componente experimental computacional. Temática y unidad de análisis: La unidad de análisis son la detección de cuerpos de agua en la provincia de Cangallo y la temática u objeto es el procesamiento de imágenes satelitales utilizando un modelo de Deep Learning. Delimitación espacial y temporal: Cuerpos de agua en la provincia de Cangallo año 2024.	V. Independiente Modelo de Deep Learning.	Arquitectura del modelo	- Tipo de red (U-Net)	Revisión del modelo.	Nominal.	Tipo.
					Entrenamiento	- Número de épocas (epochs). - Función de pérdida (Loss).	- Registro de entrenamiento. - Software (TensorFlow).	Razón.	-Número. -Valor.
					Configuración	- Número de capas. - Taza de aprendizaje (learning rate)	- Análisis del modelo. - Configuración del modelo	Razón.	-Número. -Valor.
					Procesamiento de datos	- Aplicación de Data Augmentation	- Registro d.e proceso	Nominal.	Si /No
Problemas Especificos:	Objetivos Especificos:	Hipótesis Especificas:	Diseño no experimental: No se manipula el fenómeno real, sino que se analiza imágenes satelitales existentes. Incluye un componente experimental computacional. Temática y unidad de análisis: La unidad de análisis son la detección de cuerpos de agua en la provincia de Cangallo y la temática u objeto es el procesamiento de imágenes satelitales utilizando un modelo de Deep Learning. Delimitación espacial y temporal: Cuerpos de agua en la provincia de Cangallo año 2024.	V. Dependiente Detección de cuerpos de agua.	Precisión global	- Accuracy (exactitud).	- Matriz de confusión.	Razón.	%
1. ¿Cuál es la precisión del modelo U-Net en la detección de cuerpos de agua mediante imágenes satelitales?	1. Determinar la precisión del modelo U-Net en la detección de cuerpos de agua.	1. El modelo U-Net alcanza una precisión superior al 85% en la detección de cuerpos de agua.			Desempeño de Segmentación	- IoU (Intersection over Union). - Precision (precisión positiva). - Recall (sensibilidad).	- Evaluación del modelo. - Matriz de confusión. - Matriz de confusión.	Razón.	%
2. ¿Qué nivel de desempeño presenta el modelo U-Net en la segmentación de cuerpos de agua según métricas IoU (Intersection over Union) y Accuracy (exactitud)?	2. Evaluar el desempeño del modelo U-Net en la segmentación de cuerpos de agua mediante métricas IoU y Accuracy.	2. El modelo U-Net alcanza un índice IoU superior a 0.85, evidenciando alta capacidad en la segmentación.			Identificación	- Número de cuerpos detectados correctamente	- Análisis espacial (SIG)	Razón.	Cantidad
3. ¿Existe una diferencia significativa entre la detección de cuerpos de agua mediante el modelo U-Net y el índice espectral MNDWI (Modified Normalized Difference Water Index)?	3. Comparar estadísticamente la precision del modelo U-Net frente al índice MNDWI mediante la prueba de McNemar.	3. Existe una diferencia estadísticamente significativa ($p < 0.05$) a favor del modelo U-Net frente al índice MNDWI en la detección de cuerpos de agua.			Comparación	- Diferencia con índice MNDWI.	- Prueba McNemar.	Nominal/ inferencial.	



UNSCH

FACULTAD DE
INGENIERÍA
DE MINAS, GEOLOGÍA Y CIVIL

ACTA DE SUSTENTACIÓN DE TESIS N° 32-2026-FIMGC

PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO CIVIL

En la Universidad Nacional de San Cristóbal de Huamanga, en la ciudad de Ayacucho, en cumplimiento a la **Resolución Decanal No 189-2026-FIMGC-D**, a los **diecinueve días del mes de junio de 2026**, siendo las **10:18 a.m.**, reunidos en el **Auditorio de la Escuela Profesional de Ingeniería Civil**, bajo la presidencia del **MSc. José Ernesto ESTRADA CARDENAS**, y los miembros: **Ing. Jaime Leonardo BENDEZÚ PRADO**, **Ing. Edward LEON PALACIOS** y **Ing. Saul Walter RETAMOZO FERNANDEZ**, actuando como secretario docente el **Ing. Daniel FERRUA CARRIÓN**, para proceder a la sustentación de tesis para optar el **Título Profesional de Ingeniero Civil**, del Bachiller en Ciencias de la Ingeniería Civil:

ALEX ELVEN CHAVEZ TITO

Quien presentó la tesis denominada:

**DETECCIÓN DE CUERPOS DE AGUA MEDIANTE EL PROCESAMIENTO DE
IMÁGENES SATELITALES UTILIZANDO UN MODELO DE DEEP LEARNING EN
LA PROVINCIA DE CANGALLO, 2024**

Los señores miembros del jurado luego de expuesta la tesis y absueltas las preguntas, deliberaron y declararon:

Aprobado con 17 (diecisiete)

Siendo las **11:40 a.m.** del día **19 de junio del 2026**, culmina el acto de sustentación de tesis, y en conformidad de lo actuado los miembros del jurado firmamos al pie del presente.

MSc. José Ernesto ESTRADA CARDENAS
Presidente

Ing. Jaime Leonardo BENDEZÚ PRADO
Miembro

Ing. Saul Walter RETAMOZO FERNANDEZ
Miembro

Ing. Edward LEON PALACIOS
Miembro - Asesor

Ing. Daniel FERRUA CARRIÓN
Secretario docente de la FIMGC

FACULTAD DE INGENIERÍA
DE MINAS Y CIVIL
Av. Independencia S/N
Ciudad Universitaria
Central Tel. 066 312510
Anexo 151



CONSTANCIA DE ORIGINALIDAD DE TRABAJO DE INVESTIGACIÓN

CONSTANCIA N° 21-2026-FIMGC/ASIH

El que suscribe; responsable verificador de originalidad de trabajos de tesis de pregrado con el software Turnitin, de la Escuela Profesional de **Ingeniería Civil** de la **Facultad de Ingeniería de Minas, Geología y Civil**; en cumplimiento a la **Resolución de Consejo Universitario N° 039-2021-UNSH-CU**, Reglamento de Originalidad de Trabajos de Investigación de la Universidad Nacional San Cristóbal de Huamanga y **Resolución Decanal N° 697-2024-FIMGC-UNSH-D**, dejo constancia de originalidad de trabajo de investigación, que el/la Sr./Srta.

Apellidos y Nombres : Alex Elven CHAVEZ TITO
Escuela Profesional : INGENIERÍA CIVIL
Título de la Tesis : DETECCIÓN DE CUERPOS DE AGUA MEDIANTE EL PROCESAMIENTO DE IMÁGENES SATELITALES UTILIZANDO UN MODELO DE DEEP LEARNING EN LA PROVINCIA DE CANGALLO, 2024
Evaluación de la Originalidad : 7 % Índice de Similitud
Identificador de la entrega : 3002310745

Por tanto, según los Artículos 12, 13 y 17 del Reglamento de Originalidad de Trabajos de Investigación, es **PROCEDENTE** otorgar la **Constancia de Originalidad** para los fines que crea conveniente.

En señal de conformidad y verificación se firma la presente constancia

Ayacucho, 23 de julio del 2026



UNIVERSIDAD NACIONAL DE
SAN CRISTÓBAL DE HUAMANGA
Facultad de Ingeniería de Minas, Geología y Civil

Mg. Ing. Alex Sander IRCANAUPA HUAMANI
Verificador de Originalidad de Trabajos de Tesis de Pregrado
Escuela de Formación Profesional de Ingeniería Civil

DETECCIÓN DE CUERPOS DE AGUA MEDIANTE EL PROCESAMIENTO DE IMÁGENES SATELITALES UTILIZANDO UN MODELO DE DEEP LEARNING EN LA PROVINCIA DE CANGALLO, 2024

por Alex Elven CHAVEZ TITO

Fecha de entrega: 23-jul-2026 01:15p. m. (UTC-0500)

Identificador de la entrega: 3002310745

Nombre del archivo: Te_Alex_Elven_Chavez_Tito.pdf (8.38M)

Total de palabras: 27025

Total de caracteres: 158181

DETECCIÓN DE CUERPOS DE AGUA MEDIANTE EL PROCESAMIENTO DE IMÁGENES SATELITALES
UTILIZANDO UN MODELO DE DEEP LEARNING EN LA PROVINCIA DE CANGALLO, 2024

INFORME DE ORIGINALIDAD

7%	7%	3%	4%
INDICE DE SIMILITUD	FUENTES DE INTERNET	PUBLICACIONES	TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	Submitted to Universidad Nacional de San Cristóbal de Huamanga	2%
	Trabajo del estudiante	
2	repositorio.unsaac.edu.pe	1%
	Fuente de Internet	
3	hdl.handle.net	<1%
	Fuente de Internet	
4	www.campustudionline.com	<1%
	Fuente de Internet	
5	Najar Mendoza, William Lucio. "Aplicación de sistemas de información geográfica para la evaluación de riesgos por fenómenos naturales a inundación fluvial en la ciudad urbana del distrito de Iñapari, Madre de Dios - 2024", Universidad Nacional del Altiplano de Puno (Peru), 2025	<1%
	Publicación	
6	Submitted to Uniminuto Virtual	<1%
	Trabajo del estudiante	
7	ir.lib.hiroshima-u.ac.jp	<1%
	Fuente de Internet	
8	Jhon Jairo Sáenz Gamboa. "Machine Learning en Resonancia Magnética. Describiendo la Fisiopatología del Dolor Lumbar", Universitat Politecnica de Valencia, 2024	<1%
	Publicación	
9	vsip.info	<1%
	Fuente de Internet	
10	Submitted to Universidad Rey Juan Carlos	<1%
	Trabajo del estudiante	
11	repositorio.unsch.edu.pe	<1%
	Fuente de Internet	
12	cris.usm.cl	<1%
	Fuente de Internet	

13	Fuente de Internet	<1 %
14	Calli Olvea, Javier. "Deep learning para la detección e identificación de estudiantes en el Instituto de Educación Superior Tecnológico Publico Espinar, 2023", Universidad Nacional del Altiplano de Puno (Peru), 2025 Publicación	<1 %
15	Submitted to Universidad Carlos III de Madrid - EUR Trabajo del estudiante	<1 %
16	Submitted to Universidad Tecnológica Indoamerica Trabajo del estudiante	<1 %
17	Submitted to Pontificia Universidad Catolica Madre y Maestra PUCMM Trabajo del estudiante	<1 %
18	Valdés González, Jorge L.. "Detección de anomalías en la validación de identidad biométrica facial: Contribución del fine-tuning a una red VGG-16.", Universidad de La Sabana (Colombia) Publicación	<1 %
19	idoc.pub Fuente de Internet	<1 %
20	Submitted to Universidad Catolica San Antonio de Murcia Trabajo del estudiante	<1 %
21	Submitted to Universidad Cesar Vallejo Trabajo del estudiante	<1 %
22	js19blog.wordpress.com Fuente de Internet	<1 %
23	1library.co Fuente de Internet	<1 %
24	ri-ng.uaq.mx Fuente de Internet	<1 %

Excluir citas
Excluir bibliografía

Activo
Activo

Excluir coincidencias

< 30 words