



Instructivo

Clasificación de coberturas de ortoimágenes para entrenamiento en el software PCI Geomática

Código IN-XX-XX
Versión 1
Vigente desde XX/11/2023

1. OBJETIVO

Proporcionar un modelo de clasificación supervisada de imágenes orientado a objetos de la mayor exactitud posible para las ortoimágenes X – X, con el software Focus Banff de la suite de PCI Geomatica, entregando como resultado los parámetros de entrada para la clasificación de imágenes y un archivo shapefile con las clasificaciones de cobertura.

2. ALCANCE

El presente instructivo describe el resultado obtenido al desarrollar un modelo de clasificación supervisada de imágenes orientado a objetos de la mayor exactitud posible para las ortoimágenes X – X, con el software Focus Banff de la suite de PCI Geomatica, entregando como resultado los parámetros de entrada para la clasificación de imágenes y un archivo shapefile con las clasificaciones de cobertura. Las clasificaciones de cobertura que se tuvieron en cuenta se muestran en la tabla 1.

Tabla 1. Tipos de cobertura en la clasificación.

| COBERTURA | DESCRIPCIÓN |
|------------------------|--|
| Cuerpos de agua | Corresponde a todas las superficies de agua naturales o artificiales en territorio, como ríos, lagos, lagunas, pozos, estanques, entre otros. |
| Bosques | Corresponde a toda la cobertura vegetal densa en territorio. |
| Pastos | Corresponde a la cobertura vegetal de baja densidad. |
| Vías | Corresponde a la malla vial del territorio. |
| Construcciones | Corresponde a todas aquellas estructuras hechas por el ser humano en territorio, como casas, edificios, complejos, entre otras instalaciones. |
| Suelo desnudo | Se refiere a aquel que no cuenta con una capa superficial que lo protege de la intemperie, por lo que está en riesgo de ser erosionado por agua, o viento. |

3. DEFINICIONES





- **Clasificación por "Random Trees":** Un algoritmo de clasificación utilizado en teledetección que se basa en la construcción de un conjunto de árboles de decisión aleatorios para asignar categorías a píxeles de una imagen.
- **Clasificación por "Support Vector Machine":** Un método de clasificación en teledetección que utiliza vectores de soporte para separar datos en diferentes categorías mediante la construcción de un hiperplano óptimo.
- **Cobertura:** Extensión geográfica o área de superficie de la Tierra que es observada o capturada por un sensor remoto, ya sea a través de imágenes satelitales, fotografías aéreas u otros dispositivos de observación.
- **Composición de Color (RGB):** La representación de una imagen mediante la combinación de tres canales de color: rojo (R), verde (G) y azul (B), que permite la visualización de diferentes características en la imagen.

- **Firma Espectral:** El patrón único de reflectancia o emisión de energía electromagnética de un objeto en función de diferentes longitudes de onda, que se utiliza para identificar y caracterizar objetos en imágenes de teledetección.
- **Homogeneidad:** La medida de la uniformidad o variabilidad de una región en una imagen de teledetección, que puede indicar la consistencia de la cobertura o la textura de la superficie.
- **Imagen:** Cobertura de malla cuyos valores de atributo son una representación numérica de un parámetro físico.
- **Índices de Cobertura:** Parámetros utilizados para cuantificar la presencia y distribución de diferentes tipos de cobertura terrestre (como vegetación, agua, suelo) en una imagen, a menudo basados en la relación entre bandas espectrales.
- **Metadato:** Información que describe la organización de los datos geoespaciales, la calidad de la información, sus referencias espaciales, sus entidades y atributos, la distribución de la información, entre otros.
- **Ortoimagen:** Mosaico de imágenes que, mediante proyección ortogonal a una superficie de referencia se le ha eliminado el desplazamiento debido a la inclinación del sensor y al relieve del terreno.
- **PCI Geomática:** PCI Geomática es un software poderoso para la teledetección y el procesamiento y análisis de imágenes aéreas y satelitales de la superficie terrestre. El principal objetivo de este software es llevar a cabo un procesamiento rápido de datos y un análisis avanzado de imágenes aéreas y satelitales.
- **Píxel:** El elemento más pequeño de una imagen digital en teledetección, que representa un punto en la superficie de la Tierra y tiene atributos como valor de reflectancia, temperatura, etc.
- **Reflectancia:** La fracción de radiación incidente que es reflejada por un objeto o superficie terrestre en una imagen de teledetección, generalmente expresada como un valor entre 0 y 1.
- **Resolución Espacial:** La capacidad de un sensor para distinguir detalles finos en una imagen, determinada por el tamaño del píxel y la capacidad de separar objetos cercanos.
- **Resolución Espectral:** La capacidad de un sensor para capturar información en diferentes bandas del espectro electromagnético, lo que permite la discriminación de características basadas en sus propiedades espectrales.
- **Resolución Radiométrica:** La capacidad de un sensor para detectar y distinguir diferentes niveles de intensidad de radiación electromagnética, que se relaciona con la profundidad de bits de la imagen.
- **Resolución Temporal:** La frecuencia con la que un sensor captura datos para una ubicación específica en la Tierra, medida en intervalos de tiempo.
- **Segmentación:** Proceso de dividir una imagen en segmentos más pequeños con características similares, como color, textura o forma. Este enfoque facilita la clasificación de áreas significativas en lugar de píxeles individuales, mejorando la interpretación y el análisis de la información contenida en una imagen.
- **Textura:** La apariencia superficial de una imagen de teledetección que describe cómo las características están dispuestas y se relacionan entre sí, y que puede ser importante para la interpretación y clasificación de la imagen.

4. DESARROLLO

Se desarrollaron diferentes pruebas en el software PCI Geomática Banff, con el fin de determinar la mejor configuración para clasificación de coberturas de un conjunto de imágenes. En la tabla 2 se consigna la información de las imágenes tratadas, todas ellas recuperadas del portal Colombia en Mapas del Instituto Geográfico Agustín Codazzi.

Tabla 2. imágenes tratadas en la clasificación.

| NOMBRE DE LA IMAGEN | SENSOR | DESCRIPCIÓN | IMAGEN |
|-----------------------|---------------------|---|---|
| TOPAGA | Jilin-JL-1KF01 | Ortoimagen, a la cual se le aplicó un proceso de balance radiométrico y edición de líneas de costura, garantizando la continuidad cromática y geométrica de los elementos. Este producto contiene información del municipio de Tópaga, departamento de Boyacá, República de Colombia. Tiene un área de 3590,38 hectáreas. Cuenta con un GSD de 50 cm, aplicable para cartografía a escala 1:5.000, los insumos fueron capturados el día 29 de enero de 2023 |  |
| VILLA URIBE | Vexcel Eagle Mark 3 | Ortoimagen compuesto de imágenes ortorectificadas a las cuales se les aplicó un proceso de balance radiométrico y edición de líneas de costura, garantizando la continuidad cromática y geométrica de los elementos. Este producto contiene información del Municipio de Montecristo, Centro poblado Villa Uribe, Departamento de Bolívar, República de Colombia. Tiene un área de 30,51 hectáreas. Cuenta con un GSD de 20 cm, aplicable para cartografía a escala 1:2.000, los insumos fueron capturados el día 05 de septiembre de 2023. |  |
| MONGUI | Jilin-JL-1KF01 | Ortoimagen compuesta de imágenes ortorectificadas, a las cuales se les aplicó un proceso de balance radiométrico y edición de líneas de costura, garantizando la continuidad cromática y geométrica de los elementos. Este producto contiene información del municipio de Monguí, departamento de Boyacá, República de Colombia. Tiene un área de 7334,72 hectáreas. Cuenta con un GSD de 50 cm, aplicable para cartografía a escala 1:5.000, los insumos fueron capturados el día 29 de enero de 2023. |  |
| REMOLINOS DE ARICUNTI | Vexcel Eagle Mark 3 | Ortoimagen compuesto de imágenes ortorectificadas, a las cuales se les aplicó un proceso de balance radiométrico y edición de líneas de costura, garantizando la continuidad cromática y geométrica de los elementos. Este producto contiene información del municipio de Milán, Centro Poblado Remolinos de Aricunti, Departamento de Caquetá, República de Colombia. Tiene un área de 14,71 hectáreas. Cuenta con un GSD de 20 cm, aplicable |  |

| | | | |
|--|--|--|--|
| | | para cartografía a escala 1:2.000, los insumos fueron capturados el día 31 de julio de 2023. | |
|--|--|--|--|

La metodología se realizó por medio de prueba y error, aplicando filtros pre-segmentación para cada imagen, los cuales se presentan en el anexo 1. Finalmente, los pasos seguidos en la metodología se describen en la tabla 3.

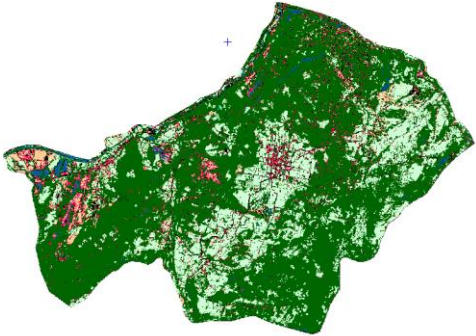
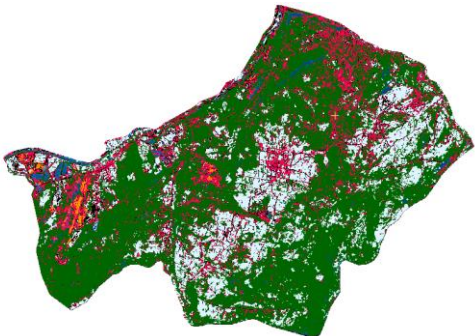
Tabla 3. Metodología general de la clasificación de coberturas en una ortoimagen.

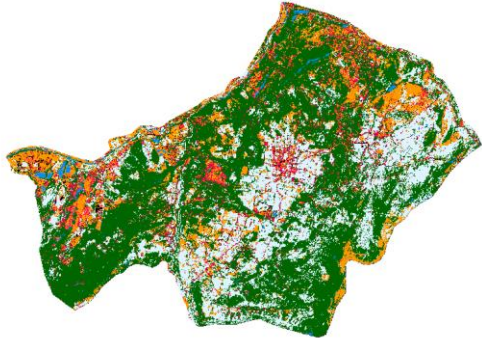

| PASO | DESCRIPCIÓN |
|------|---|
| 1 | Importar la imagen de interés al programa Geomatica Banff y posteriormente transformar dicha imagen en el formato propietario (.pix). |
| 2 | Generar y agregar filtros como nuevas capas de la imagen, previo a la segmentación. |
| 3 | Iniciar el proceso de segmentación de la imagen multibanda. |
| 4 | Calcular los atributos para entidad de la segmentación realizada, en este caso, se calculan los atributos de las siguientes categorías para todas las imágenes: Statistical (Todos) Geometrical (Todos) Vegetation Index (NDVI) |
| 5 | Realizar el entrenamiento del software, creando las clases de coberturas que se desean clasificar. En este ejercicio, se tomaron en cuenta las clases de cobertura consignadas en la tabla 1. Adicionalmente, se toman 20 muestras de entrenamiento y 5 muestras de validación para cada cobertura. |
| 6 | Seleccionar el método de clasificación supervisada, en este caso, se emplearon los métodos Random Forest y Support Vector Machine. |
| 7 | Evaluar visualmente la clasificación generada en busca de errores. |
| 8 | Evaluar estadísticamente la clasificación generada, generando un reporte con el porcentaje de precisión en la clasificación y demás parámetros de interés para evaluar el asertividad del proceso realizado. |
| 9 | Si ambas evaluaciones del modelo son favorables, se aumenta el número de muestras en el modelo de entrenamiento para mejorar las imprecisiones. En caso contrario, se descarta la clasificación y se inicia desde el paso 1 con nuevos parámetros pre y post segmentación. |

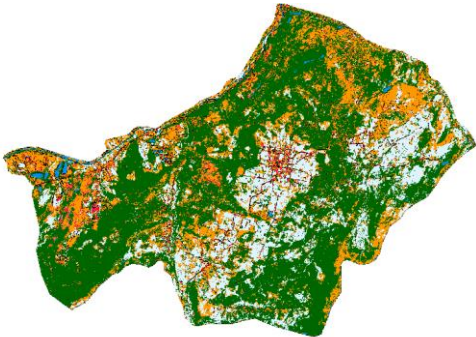
A continuación se describe el proceso específico llevado a cabo en cada imagen, condensando los principales parámetros de entrada y principales resultados de la metodología propuesta en la tabla 3 en nuevas tablas individuales

4.1. TOPAGA

A continuación se presentan las pruebas realizadas a la imagen TOPAGA, siguiendo la metodología expuesta en la tabla 3.

| | | | | | | | | |
|---|--|----|-----|-----|---------------------------------|-------|-------|---|
| 1 | Textura - Homogeneity Soil Adjusted Vegetation Index - SAVI | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 97,44 | 0,969 |  <p>Se obtienen resultados favorables en la diferenciación de pastos y vegetación, no obstante, para cuerpos de agua, vías y construcciones se mezclan entre sí algunas superficies, especialmente en las vías se requiere mejorar la precisión en la continuidad.</p> |
| 2 | Textura - Homogeneity SAVI (Soil Adjusted Vegetation Index) NDWI (Índice Diferencial de Agua Normalizado) | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 93,33 | 0,92 |  <p>Se obtienen resultados menos favorables que en el modelo anterior, teniendo en cuenta no solo el porcentaje de precisión, si no que en este caso se genera mucho ruido en la clasificación por las coberturas de suelo descubierto tomadas</p> |

| | | | | | | | | |
|---|--|----|-----|-----|---------------------------------|-------|-------|--|
| | | | | | | | | como construcciones por el modelo. Adicionalmente se presenta una mala interpretación de cuerpos de agua. |
| 3 | Textura – Homogeneity SAVI (Soil Adjusted Vegetation Index) NDWI (Índice Diferencial de Agua Normalizado) | 25 | 0,5 | 0,5 | Random Trees (RT) | 75,76 | 0,708 |  <p>Se obtiene una diferenciación más clara entre suelo desnudo y construcciones, no obstante en el recuento estadístico se obtiene un porcentaje de precisión menos favorable respecto a los casos anteriores</p> |
| 4 | Textura - Homogeneity SAVI (Soil Adjusted Vegetation Index) NDWI (Índice Diferencial de Agua Normalizado) | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 93,94 | 0,927 |  <p>Se obtiene una diferenciación aún más clara entre suelo desnudo y construcciones, obteniendo unos porcentajes de precisión más favorables. Se diferencia correctamente el suelo desnudo de pastos en gran parte de la imagen, significando una mejora del modelo anterior. No obstante, se</p> |

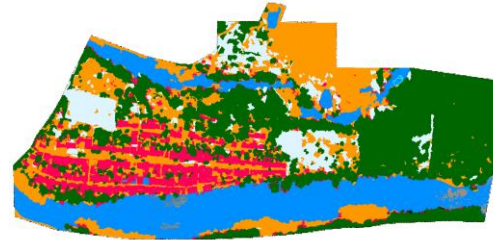
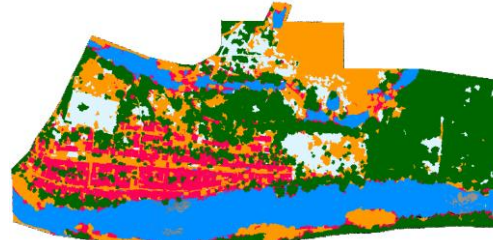
| | | | | | | | | |
|---|--|----|-----|-----|---------------------------------|-------|-------|---|
| | | | | | | | | clasifican erróneamente coberturas de vías como suelo desnudo, así como otros errores puntuales en toda la imagen. |
| 5 | Textura - Homogeneity SAVI (Soil Adjusted Vegetation Index) NDWI (Índice Diferencial de Agua Normalizado) | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 85,29 | 0,823 |  <p>Se toma la iteración anterior y se amplía en 10 unides las muestras tomadas, donde se logra una diferenciación aún más clara entre suelo desnudo y construcciones, obteniendo unos porcentajes de precisión un poco menos favorables. No obstante se diferencian correctamente suelo desnudo, pastos, y construcciones en gran parte de la imagen, significando una mejora del modelo anterior. Adicionalmente, la continuidad de las vías se ve mejorada a pesar de que se obtiene un porcentaje de precisión del 66% en su precisión individual.</p> |

Luego del trabajo realizado, se determina que la mejor configuración de clasificación para la imagen TOPAGA corresponde a la prueba numero 5, ya que presenta una adecuada diferenciación de las coberturas de suelo desnudo, construcciones y vías, que tienden a clasificarse erróneamente entre si debido a que sus firmas espectrales presentan ciertas similitudes, esto gracias a la ampliación de las muestras en el entrenamiento del modelo. Por otro lado, se obtiene un porcentaje de precisión del modelo de 85,29 %, el cual es inferior respecto a las interiores iteraciones, no obstante en la evaluación visual del modelo se corrigen errores de clasificación de todas las pruebas anteriores, teniendo una mejor continuidad y delimitación de todas las coberturas en toda la imagen, por lo que se toma como la clasificación más adecuada.

4.2. VILLA URIBE

A continuación se presentan las pruebas realizadas a la imagen VILLA URIBE, siguiendo la metodología expuesta en la tabla 3.


| | | | | | | | | |
|---|---|----|-----|-----|---------------------------------|-------|------|---|
| 1 | Textura - Homogeneity Textura - Mean | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 92,00 | 0,9 |  <p>Se obtienen resultados favorables para esta primera prueba, con una clasificación notablemente acertada, con algunos errores puntuales de continuidad en las superficies, no obstante, se logra una diferenciación precisa de todas las construcciones en el centro poblado de la imagen, pero se clasifican erróneamente superficies de suelo desnudo como construcciones en cercanías a los cuerpos de agua.</p> |
| 2 | Textura - Homogeneity Textura - Mean | 25 | 0,5 | 0,5 | Random Trees (RT) | 83,33 | 0,82 |  <p>Se obtienen resultados menos favorables que en el modelo anterior, teniendo en cuenta no solo el porcentaje de precisión, si no que en este caso se genera más ruido de construcciones en las banquetas de los cuerpos de agua.</p> |



| | | | | | | | | |
|---|---|----|-----|-----|------------------------------|-------|-------|--|
| 3 | Textura - Homogeneity* NDWI (Índice Diferencial de Agua Normalizado) Textura - Mean* | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 96,30 | 0,954 |  <p>Se obtiene una diferenciación más clara entre suelo desnudo y el espejo de agua que refleja los rayos del sol, obteniendo resultados más favorables en la continuidad de los cuerpos de agua. Además se obtiene clara diferenciación del suelo desnudo y construcciones en gran parte de la imagen.</p> |
| 4 | Textura - Homogeneity* NDWI (Índice Diferencial de Agua Normalizado) Textura - Mean* | 25 | 0,5 | 0,5 | Random Trees (RT) | 66,67 | 0,583 |  <p>Se obtiene una diferenciación buena entre suelo desnudo y construcciones, pero deficiente respecto a los modelos anteriores.</p> |

Luego del trabajo realizado, se determina que la mejor configuración de clasificación para la imagen VILLA URIBE corresponde a la prueba número 3, ya que presenta una adecuada diferenciación de las coberturas de suelo desnudo y construcciones. Por otro lado, se obtiene un porcentaje de precisión del modelo muy alto, correspondiente a 96,30 %, además de presentar una delimitación muy precisa de las construcciones las cuales por lo general son especialmente complejas en su clasificación, así como las otras coberturas.

4.3. MONGUI

A continuación se presentan las pruebas realizadas a la imagen MONGUI, siguiendo la metodología expuesta en la tabla 3.



| | | | | | | | | |
|---|--|----|-----|-----|---------------------------------|-------|-------|--|
| 1 | Textura - Homogeneity Soil Adjusted Vegetation Index - SAVI | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 82,14 | 0,777 |  <p>Se obtiene un balance positivo en el apartado de precisión, así mismo, se logra diferenciar completamente las coberturas de pastos, bosques y búsqueda. Así mismo, se evidencia que gran parte de la imagen cuenta con la cobertura de suelo desnudo.</p> |
| 2 | Textura - Homogeneity Textura - Mean | 25 | 0,5 | 0,5 | Random Trees (RT) | 53,57 | 0,423 |  <p>Se obtienen resultados menos favorables que en el modelo anterior, teniendo en cuenta no solo el porcentaje de precisión, si no que en este caso se genera más ruido de construcciones en las banquetas de los cuerpos de agua.</p> |



| | | | | | | | | |
|---|---|----|-----|-----|---------------------------------|-------|-------|--|
| 3 | Soil Adjusted Vegetation Index - SAVI NDWI (Índice Diferencial de Agua Normalizado) Textura - Mean* | 25 | 0,5 | 0,5 | Random Trees (RT) | 68,75 | 0,606 |  <p>Se obtiene un modelo que representa de manera adecuada el suelo desnudo en la escena, no obstante, se genera mucho ruido indeseado asociado a las construcciones y vías que quedan de manera generalizada en la imagen.</p> |
| 4 | Soil Adjusted Vegetation Index - SAVI NDWI (Índice Diferencial de Agua Normalizado) Textura - Mean* | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 90,63 | 0,883 |  <p>Se obtiene una diferenciación adecuada de bosques y pastos, además de una mejora considerable en el ruido de construcciones del modelo anterior, no obstante, se siguen clasificando erróneamente coberturas pequeñas de vías y suelo desnudo como construcciones. Además, se tienen espacios en coberturas oscuras de bosques en la segmentación.</p> |

Luego de las pruebas realizadas, se determina que la mejor configuración de clasificación para la imagen MONGUI corresponde a la prueba número 4, ya que presenta una adecuada diferenciación de las coberturas de suelo desnudo, pastos y vías, con algunos errores puntuales en la clasificación de vías. Por otro lado, se obtiene un porcentaje de precisión del modelo alto, correspondiente a 90.63%.

4.4. REMOLINOS DE ARICUNTI

A continuación se presentan las pruebas realizadas a la imagen REMOLINOS DE ARICUNTI, según la metodología expuesta en la tabla 3.

| | | | | | | | | |
|---|---|----|-----|-----|---------------------------------|-------|------|---|
| 1 | Textura - Homogeneity Textura - Mean | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 96,00 | 0,95 |  <p>Se obtienen resultados favorables para esta primera prueba, con una clasificación notablemente acertada, con algunos errores puntuales de continuidad en las superficies, no obstante, se logra una diferenciación precisa de todas las construcciones en el centro poblado de la imagen, pero se clasifican erróneamente superficies de cuerpos de agua y pastos.</p> |
| 2 | Textura - Homogeneity Textura - Mean | 25 | 0,5 | 0,5 | Random Trees (RT) | 36,00 | 0,2 |  <p>Se obtienen resultados deficientes respecto al modelo anterior, teniendo en cuenta no solo el porcentaje de</p> |

| | | | | | | | | |
|---|---|----|-----|-----|------------------------------|-------|-------|--|
| | | | | | | | | precisión, si no que en este caso no se logra clasificar adecuadamente los cuerpos de agua. |
| 3 | Textura - Homogeneity* NDWI (Índice Diferencial de Agua Normalizado) Textura - Mean* | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 91,67 | 0,896 |  <p>Se obtiene una peor diferenciación del suelo desnudo y construcciones en la imagen, respecto a la primera iteración a pesar de haber incluido más filtros pre-segmentación, siendo muy similar en todo lo demás.</p> |
| 4 | Textura - Homogeneity* NDWI (Índice Diferencial de Agua Normalizado) Textura - Mean* | 25 | 0,5 | 0,5 | Support Vector Machine (SVM) | 95,83 | 0,948 |  <p>Luego de mejorar el muestreo respecto a la iteración 3, se obtiene una diferenciación buena entre suelo desnudo y construcciones, así como pastos y bosques, además de una mejor continuidad en todas las coberturas</p> |

Luego de las pruebas realizadas, se determina que la mejor configuración de clasificación para la imagen REMOLINOS DE ARICUNTI corresponde a la prueba número 4, ya que presenta la mejor diferenciación de las coberturas de construcciones, cuerpos de agua, suelo desnudo, pastos y vías. Siendo la clasificación más precisa en la clasificación de las construcciones. Por otro lado, se obtiene un porcentaje de precisión del modelo alto, correspondiente a 95.83%.

4.5. OBSERVACIONES GENERALES

En términos generales, se tiene que el método de clasificación más efectivo en todas las imágenes es el Support Vector Machine (SVM), el cual otorgó resultados favorables en la evaluación visual de las imágenes y evaluación estadística. Adicionalmente, se recomienda aumentar las muestras del modelo de entrenamiento sobre una prueba que arroje resultados favorables, con el objetivo de mejorar la precisión de dicho modelo.

5. CONTROL DE CAMBIOS

Registrar las dos últimas versiones (para el caso de actualizaciones de documentos) así:

| FECHA | CAMBIO | VERSIÓN |
|----------|--|---------|
| /11/2023 | <ul style="list-style-type: none"> Se adopta como versión 1 debido a la actualización de la Cadena de Valor en Comité Institucional de Gestión y Desempeño del 3 de marzo del 2023, nuevos lineamientos frente a la generación, actualización y derogación de documentos del SGI. Hace parte del proceso de Subdirección Cartográfica y Geodésica. Se ajusta el documento según la nueva Estructura Orgánica aprobada por Decreto 846 del 29 de Julio del 2021. | 1 |

Registrar la creación del documento en versión 1 así:

| FECHA | CAMBIO | VERSIÓN |
|------------|---|---------|
| 16/11/2023 | <ul style="list-style-type: none"> Se adopta como versión 1 por corresponder a la creación del documento. Emisión Inicial Oficial. Hace parte del proceso de Subdirección Cartográfica y Geodésica. | 1 |

| Elaboró y/o Actualizó | Revisó Técnicamente | Revisó Metodológicamente | Aprobó |
|---|---|---|--|
| Nombre: Michael Rojas Cargo: Contratista | Nombre: Diego Rugeles Cargo: Contratista | Nombre: Diego Rugeles Cargo: Contratista | Nombre: Carlos Franco Prieto Cargo: Subdirector Cartografía y Geodesia. |