

Proceso de clasificación de uso y cobertura del suelo basado en secuencias temporales de datos MODIS

Carlos Matías de la Torre

6 de mayo de 2008

Índice

1. TODO	3
2. Prólogo	3
3. Introducción	3
4. Datos: Productos MODIS	4
4.1. Provisión de datos	4
4.2. Características técnicas de los productos MODIS	7
4.2.1. Reflectancia de la superficie, diaria: MOD09GQK	7
4.2.2. Calidad de la reflectancia de la superficie, diaria: MOD09GST	9
4.2.3. Índices de vegetación 16 días: MOD13Q1	9
5. Índices de vegetación	10
5.1. NDVI	12
6. Perfiles temporales y <i>características</i>	13
7. Clases	15
7.1. Homogeneidad	17
7.2. Separabilidad	17
8. Cadenas de procesamiento	19
8.1. Cadenas Originales	19
8.1.1. Reproyección y recorte	19
8.1.2. Filtrado	19
8.1.3. Creación de las secuencias temporales y de la máscara de nubes.	20
8.1.4. Cálculo de NDVI y corrección atmosférica	20
8.1.5. Cálculo y normalización de características	20
8.1.6. Clasificación	21
8.1.7. Conjunto final de clases	22
8.2. Entrenamiento	23
8.3. Cadenas mejoradas	24
8.3.1. NDVI o EVI?	25
9. Separación y Mezcla: (Split & Merge)	26
9.1. Separación	26
9.1.1. Agricultura	27
9.1.2. Las clases Florestaadas	27
9.1.3. Las pasturas	28
9.2. Mezcla	28

10. Nuevos métodos de clasificación	29
10.1. Modelo para representar imágenes y sus propiedades	30
10.2. Método basado en la distancia Mahalanobis y la distribución χ^2	31
10.2.1. Entrenamiento, definición de clases y membrecías	32
10.3. Fusión de la información	32
10.3.1. Método basado en la regla de combinación de Dempster	35
10.3.2. Método basado en la fusión mediante una Media Ponderada	36
11. Análisis y comparación de resultados	36
11.1. Matriz de confusión	36
11.1.1. Medidas de precisión	37
11.2. Resultados	39
11.2.1. Condiciones generales de las pruebas	39
11.2.2. Cadena original	39
11.2.3. Cadena mejorada	39
11.2.4. Mahalanobis y χ^2	39
11.2.5. Regla de combinación de Dempster	39
11.2.6. Promedio ponderado	39
12. Conclusión y trabajo futuro	39

1. TODO

Apéndice con criterios de calidad para el filtrado de imágenes
Traducir las clases del portugués al español y no se que del inglés al español.
Controlar nombres de las clases y de las características.
Apéndice A: con criterios de calidad para el filtrado de imágenes
Apéndice B: Explicar q no pongo código y se puede solicitar...
Apéndice C: Diagrama de bloques con el split and merge, train y classif.

2. Prólogo

El presente documento es mi trabajo final de la Licenciatura en Ciencias de la Computación de la Facultad de Matemática, Astronomía y Física de la Universidad Nacional de Córdoba. El mismo expone resultados y parte del trabajo realizado durante 6 meses a modo de pasantía en el grupo de investigación CLIME del Institute Nacional de Recherche en Informatique et en Automatique (INRIA), en Roquencourt, Francia.

3. Introducción

El grupo CLIME está involucrado en el desarrollo de modelos y herramientas para el estudio de procesos de degradación de la tierra (desertificación y deforestación) utilizando datos e imágenes satelitales.

En el mismo se ha desarrollado un sistema de procesamiento de imágenes satelitales para clasificar la cobertura terrestre. Dicho sistema se basa en el análisis de perfiles temporales de datos obtenidos a partir de series temporales de imágenes.

El sistema de procesamiento fue realizada en el marco de un proyecto, resultado de un acuerdo franco-brasileño cuyo objetivo es el de desarrollar técnicas y herramientas para apoyar las administraciones locales en la identificación y monitoreo de los procesos de degradación ambiental a través del análisis de datos de teledeteccion.

Con foco en la cuenca el Taquarí (28.000 Km²), en el Pantanal Matogrossense (Reserva Mundial de Biósfera según la UNESCO), el trabajo es también parte de un esfuerzo internacional para el desarrollo de un modelo de las interacciones entre los sistemas humanos y naturales en toda la cuenca Río de la Plata (3.000.000 Km²), que involucra a Argentina, Bolivia, Brasil, Paraguay y Uruguay.

En la actualidad, además de Brasil, CLIME esta trabajando en cooperación con Italia, Túnez y Marruecos para utilizar esta cadena de procesamiento en dichos países, donde tienen serios problemas de desertificación.

Existen diferentes cadenas de procesamiento. El funcionamiento de las mismas se basa en el análisis de índices de vegetación, los cuales son obtenidos a partir de las imágenes satelitales.

Entre los objetivos de la pasantía estaban: mejorar ciertas características fundamentales del sistema de procesamiento, implementar una nueva cadena y evaluar los resultados.

Figura 1: Cuenca del Tacuarí en el Pantanal Matogrossense, Brasil

4. Datos: Productos MODIS

Cada cadena de procesamiento utiliza un gran número de imágenes: desde 26 hasta, potencialmente, más de 300. Esto representa un gran volumen de datos: entre 12GB y 80GB.

En sus diferentes versiones, las cadenas usan distintos conjuntos de datos. Todos derivados del mismo sensor pero con características técnicas distintas. Fundamentalmente, el proceso de clasificación se realiza analizando la evolución de un índice de vegetación (capítulo 5.1), en un período de un año.

Las distintas cadenas de procesamiento pueden dividirse en dos grupos: originales y mejoradas. En las del primer grupo se pre-procesan y filtran imágenes de reflectancia de la superficie terrestre, y luego se calcula el índice de vegetación. En las cadenas “mejoradas” se utilizan imágenes en donde la mayor parte del pre-procesamiento y el índice de vegetación ya han sido calculados.

En este capítulo se explica de dónde se obtienen las imágenes y también las características técnicas de los distintos conjuntos de datos utilizados.

4.1. Provisión de datos

Las misiones principales de la NASA (United States of America’s National Aeronautics and Space Administration) se dividen en cuatro administraciones: Investigación Aeronáutica, Sistemas de Exploración, Ciencia y Operaciones Espaciales.

El Directorio de Ciencia involucra a la comunidad científica, financia investigaciones y desarrolla artefactos espaciales para responder preguntas fundamentales que

Figura 2: Cuenca del Río de la Plata

requieran una visión desde y hacia el espacio.

Uno de sus mayores componentes es el Sistema de Observación de la Tierra (o EOS por sus siglas en inglés: Earth Observing System).

El EOS tiene tres componentes principales: una serie de satélites de observación de la Tierra, un avanzado sistema de datos y equipos de científicos quienes estudian los datos. Es el primer sistema de observación en ofrecer mediciones integradas de procesos terrestres, basadas en observaciones globales de largo plazo, de la superficie terrestre, biósfera, Tierra sólida, atmósfera y océanos.

El conjunto de instrumentos a bordo de los satélites del EOS se complementan para producir las observaciones deseadas. Los datos obtenidos de diferentes aparatos se combinan, ya sea para refinar mediciones de otros o para crear productos derivados de más de un instrumento.

Los productos del EOS se clasifican de acuerdo a su nivel de procesamiento:

Nivel 0 - Datos del instrumental o carga útil, reconstruidos, no procesados, a máxima resolución; todos los artefactos de comunicación removidos (Frames de sincronización, encabezados de comunicación, etc).

Nivel 1A - Datos del instrumental, reconstruidos, no procesados, a máxima resolución, con información temporal e información auxiliar, incluyendo coeficientes de calibración radiométrica y geométrica y parámetros de georeferencia computados y agregados, pero no aplicados, a los datos de Nivel 0.

Nivel 1B - Datos de Nivel 1A procesados a unidades dependientes del sensor (no todos los instrumentos tienen un equivalente 1B).

Nivel 2 - Variables geofísicas derivadas a la misma resolución y ubicación que la fuente de datos de Nivel 1.

Nivel 3 - Variables mapeadas en grillas uniformes, con escalas de espacio-tiempo, usualmente con algún requisito de completitud y consistencia.

Nivel 4 - Resultados de modelos o de análisis de datos de más bajo nivel (por ejemplo, variables derivadas de varias mediciones).

El satélite Terra (antiguamente llamado EOS AM-1) es la primer plataforma del EOS y provee datos globales del estado de la atmósfera, tierra y océanos, así como también sus interacciones con la radiación solar y entre ellos mismos.

Uno de los instrumentos clave a bordo del satélite Terra es el Espectrorradiómetro de Formación de Imágenes de Resolución Moderada o MODIS (Moderate Resolution Imaging Spectroradiometer). Los instrumentos del MODIS son de los mejores en la ingeniería de hardware para vuelos espaciales de teleobservación.

Terra/MODIS miran toda la superficie de la Tierra cada 1 o 2 días, adquiriendo datos en 36 bandas espectrales.

[TODO: Info de las bandas MODIS?]

MODIS provee productos específicos, entre los que se incluyen:

- Temperatura de la superficie a 1Km de resolución, día y noche, con precisión absoluta de 0.2K para los océanos y 1K para la tierra.
- Color del océano.
- Fluorescencia de clorofila.
- Concentración de clorofila.
- Producción oceánica primaria, diaria y anual.
- Condiciones y productividad de la vegetación/cobertura:
- Producción primaria neta, índice de área foliar y radiación interceptada fotosintéticamente activa.
- Tipo de cobertura terrestre con identificación de cambios e identificación.
- Índices de vegetación con sensibilidad mejorada.
- Cobertura de nieve y hielo.
- Reflectancia de la superficie (con corrección de atmósfera) y albedo.
- Cobertura de nubes con 250m de resolución diurna y 1Km de resolución nocturna.

- Propiedades de las nubes (tamaño de gota, espesor óptico, presión de tope de nubes, emisividad, etc)
- Propiedades de los aerosoles (Espesor óptico, tamaño de partícula, etc.)
- Ocurrencia, tamaño y temperatura de incendios.
- Distribución global de agua precipitable,
- Etc.

La relación entre los niveles de procesamiento y los diferentes productos de tierra del MODIS puede verse en la figura 4.1.

Las distintas versiones de la cadena de procesamiento están basadas en tres productos MODIS de tierra, estos son MOD09GQK, MOD09GST (Nivel 2) y MOD13Q1 (Nivel 3). A continuación, se detallarán las características de cada conjunto de datos.

4.2. Características técnicas de los productos MODIS

Como ya se ha mencionado, las cadenas del sistema de procesamiento pueden dividirse en dos grupos: originales y mejoradas. Una de sus diferencias es el tipo de imágenes que procesan. Las cadenas en el grupo “original” trabajan con los productos MOD09GQK y MOD09GST, que son de Nivel 2. A partir de ellos se computa el índice de vegetación. Las cadenas de la serie “mejorada” trabajan con el producto MOD13Q1, el cual es de Nivel 3 y provee el índice de vegetación ya calculado.

Las ventajas y desventajas del uso de cada conjunto de datos se discute en el capítulo 8.

Veamos las descripciones de estos productos:

4.2.1. Reflectancia de la superficie, diaria: MOD09GQK

Características

(MODIS/Terra Surface Reflectance Daily L2G Global 250m SIN Grid)

Área	= $\sim 10^\circ \times 10^\circ \text{lat/long}$
Dimensiones de la imagen	= 2 (4800x4800 filas/columnas)
Tamaño promedio de archivo	= 262 MB
Resolución	= ~ 250 metros
Proyección	= Sinusoidal
Formato	= HDF-EOS
Conjuntos de Datos Científicos	= 5 (2 bandas más 3 niveles de información de calidad)

(Science Data Sets, SDSs)

Figura 3: Diagrama de los productos MODIS de tierra

Descripción

Los productos MODIS de reflectancia de la superficie se computan a partir de las bandas 1 a 7 para proveer una estimación de la reflectancia espectral de la superficie para cada banda, como sería medida a nivel del suelo en ausencia de dispersión o absorción atmosférica. Un esquema de corrección identifica gases atmosféricos, aerosoles y nubes cirrus delgadas.

Este producto contiene datos de las dos bandas cuya resolución nativa es de 250m (bandas 1 y 2). La información de calidad de este producto es provista en tres niveles de detalle: para pixels individuales, para cada banda y cada resolución, y para todo el archivo. Los datos de reflectancia de la superficie de Nivel 2 sirven de entrada para la generación de diversos productos de tierra: Reflectancia de la superficie cada 8 días, índices de vegetación, Anomalía Térmica, Hielo/Nieve, etc.

4.2.2. Calidad de la reflectancia de la superficie, diaria: MOD09GST

Características

(MODIS/Terra Surface Reflectance Quality Daily L2G Global 1km SIN Grid)	
Área	= $\sim 10^\circ \times 10^\circ$ lat/long
Dimensiones de la imagen	= 2 (1200x1200 filas/columnas)
Tamaño promedio de archivo	= 9 MB
Resolución	= ~ 1000 metros
Proyección	= Sinusoidal
Formato	= HDF-EOS
Conjuntos de Datos Científicos	= 3

Descripción

El producto MOD09GST es una versión re-estructurada de los datos de calidad existentes en el producto anteriormente descrito (MO09GQK). Resume la calidad de los productos MOD09, incluyendo correcciones atmosféricas y otras. El producto contiene información pertinente a nubes y sombra de nubes, designación de tierra y agua, aerosoles, y la fuente de datos de las correcciones realizadas en el archivo.

4.2.3. Índices de vegetación 16 días: MOD13Q1

Características

(MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid)

Área	= $\sim 10^\circ \times 10^\circ \text{lat/long}$
Dimensiones de la imagen	= 2 (4800x4800 filas/columnas)
Tamaño promedio de archivo	= 507 MB
Resolución	= ~ 250 metros
Proyección	= Sinusoidal
Formato	= HDF-EOS
Conjuntos de Datos Científicos	= 11

Descripción

Los productos de índices de Vegetación (IVs) de MODIS usan como entrada, los productos de reflectancia MODIS/Terra (MOD09) corregidos por dispersión molecular, absorción de ozono y aerosoles.

Dos IVs se producen globalmente. Uno es el IV de Diferencia Normalizada (NDVI: Normalized Difference Vegetation Index) estándar. El otro es el IV Mejorado (EVI: Enhanced Vegetation Index) con sensibilidad mejorada en las regiones de alta biomasa y una mejora en el monitoreo de la vegetación a través de una disociación de la señal del fondo de las copas y una reducción de las influencias atmosféricas. Ambos IVs se complementan en los estudios globales de vegetación y facilitan la extracción de parámetros biofísicos de la cubierta vegetal.

Los IV incluyen marcas de calidad con datos estadísticos que indican la calidad del producto y de los datos de entrada.

5. Índices de vegetación

Una de las maneras más comunes de obtener información a partir de datos satelitales es mediante el cálculo de índices. Los mismos buscan resaltar alguna característica o variable de interés (vegetación, agua, minerales, etc) mediante operaciones entre bandas.

Entre las metodologías de más amplio uso se encuentra el cociente de bandas, ya que permite separar aquellas variables con respuestas espectrales diferentes en dos o más bandas.

En particular el cálculo de **índices de vegetación** se ha desarrollado y utilizado con importantes resultados. Estos pretenden proveer un criterio para decidir si un pixel determinado representa una zona con vegetación. Algunos son más sensibles a los efectos del suelo, otros a los de la atmósfera, al estado hídrico de las plantas o a zonas quemadas.

Los índices de vegetación se basan en el particular comportamiento reflectivo de las plantas en las zonas espectrales del visible, más particularmente del rojo (ρ_R , de 0.6 a 0.7μ), y del infrarojo cercano (ρ_{IR} , de 0.7 a 1.1μ).

Las plantas en buen estado hídrico y sanitario absorben la radiación solar en el espectro visible y la utilizan como una fuente energía para la fotosíntesis. Al mismo tiempo sus células han evolucionado para reflejar y transmitir la radiación en la franja

Figura 4: Firmas espectrales de diversas coberturas

del infrarrojo cercano (que contiene aproximadamente la mitad de la energía proveniente del sol). Esto se debe a que una gran absorción de este tipo de radiación recalentaría la planta, dañando sus tejidos. Ver figura 4.

Por eso, las plantas aparecen relativamente claras en las imágenes del visible y más oscuras en el infrarrojo cercano (ver figura 5).

Figura 5: Imágenes en escala de grises. Reflectividad en el rojo (izquierda) e infrarrojo (derecha). Las zonas más oscuras en la imagen de la derecha representan vegetación.

El presente trabajo no pretende ser un catálogo de índices de vegetación, sino que solo se mencionarán algunos a modo de ejemplo.

Diferencia simple = $\rho_{IR} - \rho_R$

Cociente simple = ρ_{IR}/ρ_R

Cociente normalizado

$$NDVI = (\rho_{IR} - \rho_R)/(\rho_{IR} + \rho_R)$$

Efecto del suelo

$$SAVI = \frac{\rho_{IR} - \rho_R}{\rho_{IR} + \rho_R + L}(1 + L)$$

Efecto de la atmósfera

$$GEMI = \eta * (1 - 0,25\eta) - \frac{\rho_R - 0,125}{1 - \rho_R}$$

donde

$$\eta = \frac{2 * (\rho_{IR}^2 - \rho_R^2) + (1,5 * \rho_{IR}) + (0,5 * \rho_R)}{0,5 + \rho_R + \rho_{IR}}$$

Vegetación en áreas con contaminación atmosférica (humo, cenizas, polución)

Con datos del Thematic Mapper (TM) a bordo del satélite Landsat:

$$AFRI2,1 = (TM4 - 0,5 * TM7) / (TM4 + 0,5 * TM7)$$

Áreas quemadas

$$NDII = (\rho_{NIR} - \rho_{SWIR}) / (\rho_{NIR} + \rho_{SWIR})$$

El seguimiento temporal de estos índices aporta, entre otras cosas, a:

- Identificar patrones de humedad y precipitaciones.
- Productividad de la vegetación.
- Monitoreo de variables como el área foliar.
- Estudiar los efectos del calentamiento global sobre la vegetación.
- Predicción de cosechas.
- Identificar condiciones de sequía.
- Evaluar riesgos de incendio.

Según Chuvieco [1], para aplicar estos índices con rigor deberán aplicarse previamente las correcciones atmosféricas y la conversión de ND (numero digital, es el valor que devuelve el sensor) a reflectividades. De todas maneras, siempre que no se pretenda conceder un valor físico a los resultados, el índice puede aplicarse directamente a los ND originales de la imagen. En este caso la valoración será relativa, pero sigue manteniéndose el mismo principio: cuanto mayor sea el resultado obtenido, tanto mayor será el vigor vegetal presente en la zona observada.

5.1. NDVI

Las cadenas de procesamiento mencionadas en este proyecto trabajan con el índice de vegetación de diferencia normalizado: NDVI (por sus siglas en inglés, *Normalized Difference Vegetation Index*). El mismo es un cociente entre las bandas del rojo y del infrarrojo cercano:

$$NDVI = \frac{\rho_{IR} - \rho_R}{\rho_{IR} + \rho_R} \quad (1)$$

Una de sus ventajas principales, además de su simplicidad, es que toma valores dentro de márgenes conocidos $(-1, 1)$, lo que hace más fácil su interpretación.

Se utiliza principalmente para distinguir vegetación sana de otras cubiertas. Además de lo mencionado anteriormente, permite:

- Identificación de áreas Forestadas-deforestadas;
- Evaluación del estado de la vegetación y su grado de estrés hídrico;
- Separación entre distintos tipos de masas vegetales;
- Monitoreo de plagas;

Aparte existen numerosas variables que se pueden derivar de este índice como por ejemplo: contenido de agua en las hojas, productividad neta de la vegetación, contenido de clorofila en la hoja, dinámica fenológica, evapotranspiración potencial, etc.

En principio, las cadenas funcionarían con otros índices. Uno de los que se probó fue el Índice de Vegetación Mejorado o **EVI** (por sus siglas en inglés, *Enhanced Vegetation Index*) ya que es provisto como un producto MODIS (al igual que el NDVI, ver capítulo 4). El EVI es mejorado en el sentido que incluye factores de corrección tendientes a atenuar los efectos del suelo y de la dispersión atmosférica.

$$EVI = \frac{\rho_{IR} - \rho_R}{\rho_{IR} + C_1 * \rho_R - C_2 * \rho_{Blue} + L} \quad (2)$$

Donde:

- L es un factor de corrección del suelo y
- C_1 y C_2 intentan corregir efectos de la dispersión atmosférica.

Fuente:

Autor: Matías Hernán Parimbelli

Universidad CAECE - Técnicas Espaciales de Análisis Tutorial creado para la Cátedra Técnicas Espaciales de Análisis de la Licenciatura en Gestión Ambiental- Universidad CAECE

[1] [http://caece.edu.ar/tea/Tutoriales/Tutorial%20Multispec.](http://caece.edu.ar/tea/Tutoriales/Tutorial%20Multispec.%20Calculo%20de%20NDVI.pdf)

[%20Calculo%20de%20NDVI.pdf](http://caece.edu.ar/tea/Tutoriales/Tutorial%20Multispec.%20Calculo%20de%20NDVI.pdf)

[2] geoagris01.com.ar/_aerodreams/indices-vegetacion.pdf

6. Perfiles temporales y *características*

La superficie terrestre tiene distintas cubiertas: ciudades, cultivos, Florestas, ríos, océanos, etc. Cada una de ellas puede clasificarse en función de la vegetación que posee y del comportamiento de la misma con el paso del tiempo. Por ejemplo, un Floresta tiene una gran cantidad de vegetación durante todo el año, mientras que un río es el caso opuesto.

Si para la misma zona geográfica se calcula el NDVI durante un período determinado, a intervalos regulares de tiempo, para cada pixel queda definido un perfil temporal que representa la evolución de la vegetación durante dicho período.

Figura 6: Ejemplo de un perfil temporal de NDVI.

El proceso de clasificación realizado por CLIME se basa en el análisis de dichos perfiles temporales. El comportamiento natural de la vegetación en los distintos tipos de cobertura de la superficie terrestre, se manifiesta como propiedades o “*características*” estadísticas del perfil temporal. A modo de ejemplo podemos mencionar que para los pixels con Florestas el NDVI promedio será alto, mientras que en las montañas o el desierto, el promedio será muy bajo.

Luego, uno de los componentes principales de la cadena son las denominadas *características* (features, en inglés).

En la práctica, a pesar del pre-procesamiento aplicado a las imágenes, los perfiles son ruidosos. Luego, para el cálculo de algunas *características*, es necesario que el perfil sea suavizado. Por eso, para cada pixel, se aproxima el perfil temporal con un polinomio.

Figura 7: Ejemplo de un perfil temporal de NDVI y su polinomio de ajuste.

Luego, un perfil temporal de NDVI se modela a partir de las siguientes *características*:

Id	Nombre	Descripción
1	Media	Promedio
2	FechaMax	Fecha del máximo valor de NDVI en el perfil.
3	Delta45Dias	Promedio del NDVI de los últimos 45 días - promedio del NDVI de los primeros 45 días
4	ValAbsPendiente	Media de los valores absolutos de la derivada de la curva, en cada fecha.
5	nModos	Número de modos
6	Min	Mínimo valor de NDVI
7	DeltaMaxMin	Máximo valor de NDVI - Mínimo valor de NDVI
8	StdDev	Desvío estándar de los valores de NDVI
9	Modo - fechaPico	Fecha del pico del modo principal
10	Modo - duración	Días desde el inicio al final del modo principal
11	Modo - amplitud	Medida de la altura del modo principal.
12	Modo - pendienteDerecha	Cantidad de Días entre el pico del modo principal y la mitad de su amplitud, hacia la derecha.

En particular para el cálculo de nModos y sus *características* derivadas, la aproximación polinomial es fundamental.

El uso de las *características*, además de ayudar aumentando la interpretabilidad, es una conocida técnica de reducción de dimensionalidad cuyo desarrollo se escapa al alcance de este proyecto.

La selección del conjunto final de *características* no es arbitraria. además de ser las que mejor representan los perfiles temporales, deben poseer ciertas propiedades (como homogeneidad y separabilidad) que serán explicadas más adelante (ver capítulo 7).

Para mayor información sobre el cálculo y selección de *características* y marco teórico como estrategia de reducción de dimensionalidad, referirse a [1] (*Tesis de Milton*).

7. Clases

Cada cadena implementa un clasificador supervisado para discriminar distintas clases a partir de una serie de imágenes. Los clasificadores supervisados deben ser “entrenados” en el reconocimiento de cierto conjunto predefinido de clases. Luego el conjunto de clases de interés debe ser conocido de antemano y es necesario tener “muestras¹” de cada una de ellas.

Este es un tema muy importante y en el cuál se diferencian las distintas cadenas de procesamiento. Para este trabajo, los datos de campo necesarios para la validación

¹Una muestra de una clase es un conjunto significativamente grande de pixels pertenecientes a esa clase.

de los resultados y para el entrenamiento de los clasificadores provienen de la cuenca del Taquarí ² (Ver 3). Las clases definidas en la verdad de campo son:

1. Agricultura
2. Agua
3. Eucaliptus
4. Sabana tropical
5. Sabana Florestada
6. Floresta
7. Floresta Estacional
8. Vegetación Ciliar
9. Sabana tropical + Vegetación Ciliar
10. Sabana tropical + Floresta
11. Floresta + Sabana tropical
12. Sabana + Floresta Estacional
13. Pasturas
14. Urbana

Cuadro 1: Listado de clases originales

Estas clases fueron provistas en una imagen Landsat por el gobierno de Brasil.

Luego, el objetivo principal del sistema es el de identificar la mayor cantidad de clases de ese conjunto, en el Área de interés.

Pero existen serias restricciones impuestas por los datos que no permiten discriminar todas esas clases (Ver 8.1). Teniendo en cuenta el capítulo 6, los pixels de una misma clase deberían tener similares *características* y, al mismo tiempo, pixels de clases diferentes deberían tener *características* diferentes. Esas propiedades se denominan **Homogeneidad** y **Separabilidad** y se formalizan de la siguiente manera:

²N. del A.: Durante el transcurso de la pasantía se iniciaron las pruebas para utilizar el sistema con datos de Túnez y del sur de Italia.

Figura 8: Imagen Landsat provista por el Gobierno de Brasil con las 14 clases originales. Usada como verdad de campo.

7.1. Homogeneidad

Hay muchas formas de interpretar la homogeneidad. Básicamente decimos que una clase es homogénea si los pixels de dicha clase evolucionan de manera similar con el tiempo.

Esto se traduce en que los perfiles temporales de los pixels de una misma clase son similares.

Luego, sea C una clase de tamaño n . Sean x_1^d, \dots, x_n^d los valores de los pixels de C en la fecha d .

Entonces, sea σ_C^d el desvío estándar de la clase C en la fecha d . Sea N la cantidad de imágenes (equivalentemente, la cantidad de días). Luego, la homogeneidad de la clase C se define como:

$$\sigma_C = \frac{1}{N} \sum_{d=1}^N \sigma_C^d \quad (3)$$

Luego, un valor bajo de σ_C indica que la clase es homogénea.

7.2. Separabilidad

Así como una misma clase debe ser homogénea, es deseable que clases diferentes se comporten diferente. Luego, los perfiles temporales de pixels de distintas clases deberían ser diferentes.

El tema de la separabilidad de clases es uno muy tratado en el ámbito de la teledetección. Entre las medidas más utilizadas para determinar la separabilidad entre

dos clases, c y s , cabe nombrar *Jeffries-Matusita* (JM) y *Divergencia Transformada* (*Transformed Divergence*) (TD):

Sean $p_c(x)$ y $p_s(x)$ las densidades de las clases c y s , X el espacio de las *características*. Luego se define

$$JM(c, s) = \int_X \left[\sqrt{p_c(x)} - \sqrt{p_s(x)} \right]^2 dx \quad (4)$$

$$TD(c, s) = 2 \left[1 - \exp \left(-\frac{D}{8} \right) \right] \quad (5)$$

Donde

$$D = \int_X [p_c(x) - p_s(x)] \ln \frac{p_c(x)}{p_s(x)} dx$$

El desarrollo y fundamento teórico de estas medidas escapan el alcance de este trabajo, para más información referirse a [1]³

A los fines prácticos de este trabajo, se definió una medida de separabilidad propia.

Sea F el número de *características* seleccionadas para calcular la separabilidad.

Sean $1 \dots N$ las clases, S_i el tamaño de la clase i , $x_1^c \dots x_{S_c}^c$ los vectores de la clase c . Luego:

$$\bar{x}_f^c = \frac{\sum_{i=1}^{S_c} x_{i,f}^c}{S_c} \quad (6)$$

es la media de la *característica* f en la clase c , para $f \in [1, F]$, $c \in [1, N]$

$$\sigma_f^c = \sqrt{\frac{\sum_{i=1}^{S_c} (x_{i,f}^c - \bar{x}_f^c)^2}{S_c - 1}} \quad (7)$$

es el desvío estándar de los valores de la *característica* f en la clase c , para $f \in [1, F]$, $c \in [1, N]$

Luego, se define la separabilidad entre las clases c y s como:

$$d(c, s) = \sum_{f=1}^F \sqrt{\frac{(\bar{x}_f^c - \bar{x}_f^s)^2}{MAX(\sigma_f^c, \sigma_f^s)}} \quad (8)$$

De esta forma, las clases con perfiles diferentes estarán separadas en el espacio de las *características* y su separabilidad será mayor. El factor $MAX(\sigma_f^c, \sigma_f^s)$ influye de tal forma que una clase no homogénea (con altos valores de desvío estándar en sus *características*) tendrá poca separabilidad en general.

³P. H. Swain y R. C. King "Two effective feature selection criteria for multispectral remote sensing", Presented at The International Joint Conference on Pattern Recognition (1973), Washington, D.C., U.S.A.

8. Cadenas de procesamiento

Como ya se ha mencionado, las cadenas pueden dividirse en *originales* y *mejoradas*. Estas se diferencian principalmente en el tipo de datos que procesan pero esto a su vez influye en toda la cadena. En el presente capítulo se describen ambos tipos.

8.1. Cadenas Originales

Las cadenas originales procesan imágenes con datos diarios de reflectancia de la superficie, en los canales rojo e infrarrojo (a partir de los cuales se computa el NDVI, ver capítulo 5). Además utilizan datos de calidad para filtrar las imágenes. Para mayor detalle, referirse al capítulo 4 y a [1: **Tesis de Milton**].

Las etapas de procesamiento en una cadena *original* son:

1. Reproyección y recorte.
2. Filtrado de imágenes de mala calidad.
3. Creación de las secuencias temporales y de la máscara de nubes.
4. Cálculo del NDVI y corrección atmosférica.
5. Cálculo y normalización de características
6. Clasificación.
7. Evaluación de los resultados.

Fuera de la cadena : Entrenamiento del clasificador.

Cuadro 2: Pasos de procesamiento de las cadenas originales

Se procederá a la descripción de cada etapa:

8.1.1. Reproyección y recorte

Las imágenes MODIS provistas por la NASA usan una proyección geográfica Sinusoidal (ver capítulo 4) y abarcan un área del globo terrestre mucho mayor que el área de estudio. El software desarrollado por el grupo CLIME trabaja con la proyección Mercator.

Por eso, las imágenes deben ser reprojectadas y recortadas para poder ser procesadas y para disminuir el volumen de datos.

8.1.2. Filtrado

Gracias al producto MOD09GST, se tienen datos de calidad para cada imagen. A partir de los mismos se definió un criterio (ver apéndice XXX) que sirve para filtrar las

imágenes de mala calidad: *contaminadas* por nubes, sombras de nubes, cirrus y altas cantidades de aerosoles.

Esto permite disminuir el volumen de datos a procesar ya que, para un año típico en la zona de estudio, el número de imágenes resultantes es de aproximadamente 71 (contra 365).

8.1.3. Creación de las secuencias temporales y de la máscara de nubes.

A partir de las reflectancias y los metadatos del producto MOD09GQK, se crea para cada fecha una máscara de nubes: es una imagen con 0 en los pixels despejados y 1 en los pixels cubiertos por nubes.

Además, se calcula el NDVI en cada fecha, para cada punto y se crea una secuencia temporal de NDVI.

8.1.4. Cálculo de NDVI y corrección atmosférica

A partir de las máscaras de nubes, se eliminan los pixels contaminados de la secuencia de NDVI y estos son corregidos aplicando una interpolación temporal.

Luego, pueden aplicarse distintos filtros a cada punto, basados en su perfil temporal (por ejemplo, filtros de mediana o de varianza).

8.1.5. Cálculo y normalización de características

Una vez obtenidos perfiles temporales de buena calidad para cada punto, se computan las características. El conjunto utilizado es:

Id	Nombre	Descripción
1	Media	Promedio
2	FechaMax	Fecha del máximo valor de NDVI en el perfil.
3	Inicio	Valor de NDVI el primer día
4	DeltaInicioFin	NDVI del último día - NDVI del primer día
5	Delta150Dias	Promedio del NDVI durante los últimos 150 días - Promedio del NDVI durante los primeros 150 días
6	ValAbsPendiente	Media de los valores absolutos de la derivada de la curva, en cada fecha.
7	nModos	Número de modos
8	Min	Mínimo valor de NDVI
9	DeltaMaxMin	Máximo valor de NDVI - mínimo valor de NDVI
10	StdDev	Desvío estándar de los valores de NDVI
11	Modo - fechaPico	Fecha del pico del modo principal
12	Modo - duración	Das desde el inicio al final del modo principal

Debido a la naturaleza del clasificador (ver 8.1.6) y a las diferencias intrínsecas de las *características*, se procede a normalizar los valores de las mismas para que posean una distribución Normal con media 0 y varianza 1.

Una vez obtenidos perfiles temporales de buena calidad para cada punto, se computan las *características*. Debido a la naturaleza del clasificador (ver 8.1.6) y a las diferencias intrínsecas de las *características*, se procede a normalizar los valores de las mismas para que posean una distribución Normal con media 0 y varianza 1.

8.1.6. Clasificación

El algoritmo de clasificación en una cadena original implementa un clasificador de *máxima verosimilitud*. Este es un abordaje puramente estadístico que asume que todas las variables (en este caso las *características*) son de origen continuo y tienen una distribución Normal o *Gaussiana* multi-dimensional, en cada clase. Estas propiedades fueron corroboradas mediante un análisis estadístico de los datos, cuya descripción escapa al desarrollo de este trabajo.

Para un pixel (representado con el vector de características), se calcula la probabilidad de que pertenezca a cada clase. Luego, se asigna dicho pixel a aquella clase con mayor probabilidad (puede dejarse el pixel no clasificado, si la máxima probabilidad no supera un umbral determinado. El tema de los pixels no clasificados no se trata en este trabajo, para mayor información referirse a [1] *tesis de Milton*).

Luego, el método se basa en la fórmula de Bayes, descrita por la ecuación 9:

$$P(w_j|x) = \frac{p(x|w_j)P(w_j)}{\sum_i p(x|w_i)P(w_i)} \quad (9)$$

Donde:

- $P(w_j)$: probabilidad *a priori* para la clase w_j
- $P(w_j|x)$: probabilidad *a posteriori* para la clase w_j dada la variable x de entrada.
- $p(x|w_j)$: función de densidad de probabilidad de la variable x , dada w_j .

Por la suposición de Normalidad de la distribución de las *características*, puede modelarse $p(x|w_j)$ como una Normal con media y matriz de covarianza estimadas a partir de los datos de la zona de entrenamiento (el entrenamiento del clasificador se trata en el capítulo 8.2).

Pero debido a la existencia de *características* de naturaleza discreta (como el número de modos), el modelo de *máxima verosimilitud* debe modificarse para adaptarse a este hecho. Luego el modelo extendido para admitir la existencia de una variable discreta es:

$$P(w_j|x) = \frac{p(x_{cont}|w_j, x_{disc})P(x_{disc}|w_j)P(w_j)}{\sum_i p(x_{cont}|w_i, x_{disc})P(x_{disc}|w_i)P(w_i)} \quad (10)$$

Donde:

- x_{cont} : Variables continuas en el conjunto de datos.

- x_{disc} : la variable de naturaleza discreta (como el *número de modos*).
- $p(x_{disc}|w_j)$: función condicional de masa de probabilidad de la variable discreta, dada una clase w_j

Para mayor información sobre este tema referirse a [1]*tesis de Milton*.

8.1.7. Conjunto final de clases

Para que el clasificador de máxima verosimilitud d buenos resultados, es necesario que las clases a discriminar sean *separables* y *homogéneas* (ver capítulo 7). Por eso no pueden identificarse todas las clases del capítulo 7.

En la imagen de abajo puede verse un ejemplo de varias clases del conjunto original que se comportan de manera similar, de tal forma que no pueden ser discriminadas.

Figura 9: Media del NDVI, por fecha, para distintas clases.

Uno de las principales falencias en las cadenas originales y mejoradas es que el análisis y selección del conjunto final de clases no fue automatizado sino que es el resultado de un proceso “manual”, realizado por los desarrolladores del sistema de procesamiento.

Luego, mediante un proceso de *separación y mezcla* (*split & merge*) que ser explicado más adelante (ver capítulo 9), el conjunto final de clases con el que trabaja la cadena original es⁴:

- Agricultura 1
- Agricultura 2
- Agricultura 3
- Floresta
- Pasturas

⁴Uno de los objetivos principales de la pasantía que da origen al presente trabajo fue el de automatizar este proceso de selección del conjunto final de clases. Ver capítulo REFERENCIA CAPITULO CADENA NUEVA

- Urbana

A continuación se muestran ejemplos de perfiles temporales típicos de cada una de estas clases:

Figura 10: Perfiles temporales típicos de cada clase principal. De izquierda a derecha, de arriba a abajo: Vaca, Perro, Gato...

8.2. Entrenamiento

De forma general, un *clasificador supervisado* es aquel que requiere de *muestras* de cada clase a discriminar, a partir de las cuales “aprende” sobre cada una de ellas.

En el caso de una cadena original, cada clase w_j se modela mediante una función ***Normal multi-dimensional***, N , en el espacio de las *características*. Para ello debemos determinar dos parámetros: la media, $\mu(w_j)$ y la matriz de covarianza $\Sigma(w_j)$ de tal forma que la función de densidad de probabilidad dada una clase w_j (ver ecuación 9) es:

$$p(x|w_j) \sim N(\mu(w_j), \Sigma(w_j)) \quad (11)$$

donde

$$N(\mu, \Sigma) = \frac{1}{(2\pi)^{1/2} \det(\Sigma)^{1/2}} e^{\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu)}$$

Luego, en la cadena original, la principal tarea en la etapa de entrenamiento es la de calcular los parámetros $\mu(w_j)$ y $\Sigma(w_j)$ para cada clase.

Para ello es necesario contar con muestras de cada clase. A partir de las mismas, se estiman los parámetros buscados. REFERENCIA SPLIT y MERGE

Lo que se hizo fue dividir la zona de interés en dos: se utiliza la zona inferior (sur) como *área de entrenamiento* y la superior (norte) como *área de prueba*.

[Imágenes máscara training and testing áreas]

De esta manera, se utilizan los pixels del área de entrenamiento para entrenar el clasificador y el área de pruebas para testarlo y verificar los resultados.

Es importante clarificar que el entrenamiento no es una etapa más dentro de la cadena de procesamiento. Una vez delimitada la zona de interés, se obtiene un área de entrenamiento a partir de la cual se calculan los parámetros necesarios para el clasificador. Luego, las clases quedan completamente modeladas y la cadena puede utilizarse cuantas veces sea necesario sin tener que volver a repetir el entrenamiento.

Es importante notar que el entrenamiento de las cadenas originales y mejoradas sigue siendo un proceso manual, basado en el análisis estadístico de las *características* y de las clases (su separabilidad y homogeneidad). Esto es uno de los puntos que se mejoran en la nueva cadena (ver capítulo REFERENCIA CAPITULO NUEVA CADENA).

8.3. Cadenas mejoradas

El esquema de procesamiento de las cadenas mejoradas es muy similar a las originales. Funcionalmente, se simplifica el preprocesamiento de las imágenes ya que se utiliza el producto MOD13Q1 (ver capítulo 4) el cual:

- Provee una imagen cada 16 días, asegurando altos niveles de calidad.
- Provee el NDVI, por lo que no hay que calcularlo.
- Viene corregido por efectos atmosféricos, de aerosoles y nubes.

Por esas razones, el software que implementa una cadena mejorada cambia (se simplifica) con respecto a una original.

Si bien se mantiene el método de clasificación (ver capítulo 8.1.6) y el conjunto final de clases, cambian un poco las *características* con que trabaja. El conjunto final de características de las cadenas mejoradas es:

Id	Nombre	Descripción
1	Media	Promedio
2	FechaMax	Fecha del máximo valor de NDVI en el perfil.
3	Delta45Dias	Promedio del NDVI de los últimos 45 días - promedio del NDVI de los primeros 45 días
4	ValAbsPendiente	Media de los valores absolutos de la derivada de la curva, en cada fecha.
5	nModos	Número de modos
6	Min	Mínimo valor de NDVI
7	DeltaMaxMin	Máximo valor de NDVI - mínimo valor de NDVI
8	StdDev	Desvío estándar de los valores de NDVI
9	Modo - fechaPico	Fecha del pico del modo principal
10	Modo - duración	Das desde el inicio al final del modo principal

Luego, el esquema de trabajo es:

1. Reproyección y recorte.
2. Creación de las secuencias temporales y de la máscara de nubes.
3. Corrección atmosférica.
4. Cálculo y normalización de características
5. Clasificación.

8.3.1. NDVI o EVI?

El producto MOD13Q1 con el que trabajan las cadenas mejoradas, provee otro índice de vegetación además del NDVI: el **EVI** (*Enhanced Vegetation Index*).

Este índice *mejorado* corrige el efecto de la señal de fondo en zonas de alta biomasa y algunos efectos atmosféricos. Se define como:

$$EVI = 2 \frac{(\rho_{nir} - \rho_{red})}{(L + \rho_{nir} + C_1 \rho_{red} + C_2 \rho_{blue})} \quad (12)$$

donde

- ρ son reflectancias,
- L es un termino de ajuste para la señal de fondo,
- C_1 y C_2 son valores de ajuste para el canal azul usado en la corrección de aerosoles para el canal rojo.

Si bien puede usarse la cadena utilizando el EVI en vez del NDVI, en la práctica se obtienen mejores resultados con el segundo. De todas formas, el EVI sigue siendo una alternativa y es particularmente prometedor en zonas con menos vegetación.

9. Separación y Mezcla: (Split & Merge)

Como ya se ha explicado en capítulos anteriores es importante que el conjunto final de clases posea ciertas propiedades de homogeneidad y separabilidad. Para lograr esto, en las cadenas existentes se realizaba un trabajo previo de análisis de los datos. El usuario a este nivel definía el conjunto final de clases proveiendo muestras de cada una, para el entrenamiento.

Uno de los objetivos para este trabajo fue el de automatizar ese proceso o definir criterios para facilitarlos. Esto es extremadamente complicado si se tienen en cuenta todos los posibles escenarios de aplicación del sistema: las clases de interés en una zona como el Amazonas serán necesariamente distintas a aquellas clases que pueden definirse en una región semi-desértica como el sur de Italia.

Inicialmente la metodología de *separación y mezcla* (split and merge) que se desarrolló propone el siguiente procedimiento:

1. Separación: dividir cada clase en sub-clases de tal forma que estas sean homogéneas.
2. Mezcla: crear una jerarquía con todas las sub-clases en la base. En cada nivel se unen las dos clases con menor separabilidad. Finalmente, en el nivel superior quedará solo una clase (unión de todas las otras).

Debido a un profundo conocimiento de las clases y luego de muchos análisis que escapan alcance de este trabajo, de encontró que esa metodología no puede ser aplicada tal cual. Debido a las diferencias entre las clases, no tiene sentido, por ejemplo, intentar mezclar sub-clases de agricultura con sub-clases de Floresta.

Además, aplicar el mismo método de separación para clases tan diversas no es conveniente, ni hace falta mezclar todas las clases.

Finalmente, la metodología de separación y mezcla implementada fue la siguiente:

9.1. Separación

Primeramente, para realizar la separación se investigaron diversos algoritmos de colesteroína⁵ pero finalmente, cambiando generalidad por desempeño, se optó por utilizar un método más especializado, basado en el conocimiento de las clases y las *características*.

Entonces, las clases se separaran en grupos:

- la agricultura,
- las clases Forestadas: Sabana tropical, Sabana Florestada, Floresta, Floresta Estacional, Vegetación Ciliar, Sabana tropical + Vegetación Ciliar, Sabana tropical + Floresta, Floresta + Sabana tropical, Sabana + Floresta Estacional,

⁵ver Añil K. Jaén and otear, Testículo Paterna Precognición: A Reviene, IEEE Transacciones no paterna análisis and machín inteligencia, vil. 22, No. 1, Jaguar 2000

- las Pasturas.

Algunas clases (como el agua y las zonas urbanas) no son separadas.

9.1.1. Agricultura

La agricultura se separa de acuerdo al número de Modos. Cada modo en el perfil temporal de un pixel de esta clase representa un ciclo de siembra-cosecha.

Después, pixels con solo un modo son separados en dos, de acuerdo a la fecha del pico del modo. De esta manera se diferencian cultivos en diferentes épocas del año: cultivos tempranos y cultivos tardíos.

Figura 11: Distintos perfiles temporales de pixels pertenecientes a agricultura.

9.1.2. Las clases Florestadas

La separación de las clases Florestadas no se hace para identificar nuevas florestas, sino para identificar la **deforestación**. El perfil temporal de un pixel perteneciente a una clase Florestaada se caracteriza por mantener altos valores de NDVI durante todo el año. Pero el perfil de un área deforestada se parecerá a una clase Florestaada al principio y luego presentará un claro desnivel en el perfil temporal, a partir del cual mantendrá bajo (más similar a las pasturas). En la figura 9.1.2 pueden verse los perfiles temporales de dos pixels: el de la izquierda representa una clase Florestaada y el de la derecha representa la deforestación.

Para identificar esta cualidad, se utilizan las *características **DeltaMaxMin** y **Delta45Dias***. En la figura 9.1.2 se pueden ver los valores de dichas características para los pixels pertenecientes a la clase 10. Principalmente, se caracterizan por bajos valores de ***DeltaMaxMin*** y valores de ***Delta45Dias*** cercanos a 0. Esto se debe a que, para la misma época, cada año, el nivel de biomasa será similar.

Pero en el recuadro rojo podemos ver un conjunto de pixels con altos valores en ambas *características*. Estos representan las áreas deforestadas. La separación se aplica a *cada una de estas clases por separado*.

Figura 12: Perfiles temporales de NDVI: el de la izquierda representa una clase Florestaada y el de la derecha deforestación.

Figura 13: Valores de los pixels de la clase 10 para dos de sus *características*: en el eje horizontal *DeltaMaxMin* y en el eje vertical *Delta45Dias*

9.1.3. Las pasturas

Pasturas es la clase más extensa. Se caracteriza por un bajo nivel de NDVI, con poca variación durante el año. Pero en ocasiones se detectan modos en el perfil temporal de estos pixels, por lo que se decidió dividir esta clase de acuerdo con el número de modos para lograr mayor separabilidad.

9.2. Mezcla

La mezcla solo se aplica en el grupo de las clases Florestadas. Se busca identificar la mayor cantidad de clases Florestadas y deforestadas como sea posible. La metodología aplicada es:

1. Inicio: Sub-clases Florestadas
2. Si hay más de una clase:
3. Sabana tropical
 - a) Computar la *separabilidad* entre todos los pares de clases.
 - b) *Mezclar* las dos clases con menor separabilidad.

c) goto 2

La medida de *separabilidad* utilizada se introdujo en la ecuación 8 (capítulo 7), esto es:

Sea F el número de *características* seleccionadas para calcular la separabilidad.

Sean $1 \dots N$ las clases, S_c el tamaño de la clase c , $x_1^c \dots x_{S_c}^c$ los vectores de la clase c . Luego:

$$\bar{x}_f^c = \frac{\sum_{i=1}^{S_c} x_{i,f}^c}{S_c} \quad (13)$$

es la media de la *característica* f en la clase c , para $f \in [1, F]$, $c \in [1, N]$

$$\sigma_f^c = \sqrt{\frac{\sum_{i=1}^{S_c} (x_{i,f}^c - \bar{x}_f^c)^2}{S_c - 1}} \quad (14)$$

es el desvío estándar de los valores de la *característica* f en la clase c , para $f \in [1, F]$, $c \in [1, N]$

Luego

$$d(c, s) = \sum_{f=1}^F \sqrt{\frac{(\bar{x}_f^c - \bar{x}_f^s)^2}{MAX(\sigma_f^c, \sigma_f^s)}}$$

es la separabilidad entre las clases c y s .

Luego, el método propuesto no es más que un algoritmo jerárquico. En cada nivel de la jerarquía se define un conjunto de clases tal que, más abajo están todas las clases iniciales y en la cima solo hay una.

Una forma de definir un conjunto final de clases es escogiendo un límite, t , tal que el conjunto final de clases es el primer nivel de la jerarquía en que, para cada par de clases, la separabilidad es siempre mayor que t .

10. Nuevos métodos de clasificación

Uno de los principales objetivos de este trabajo es el de definir e implementar nuevas metodologías de clasificación, para reemplazar el algoritmo de Máxima Verosimilitud existente (ver capítulo 8.1.6).

A diferencia del anterior, basado en la teoría de las probabilidades, los nuevos métodos se desarrollaron dentro del ámbito del **pattern recognition**. Debe quedar claro que no se investigó en profundidad esta rama de la informática, sino que se incurrió de forma superficial, explorando nuevas y posibles metodologías de clasificación.

Finalmente, se rediseñaron las fases de entrenamiento y clasificación (Ver el listado 2 en el capítulo 8) y se implementaron 3 procedimientos distintos.

En la figura 14 podemos ver un diagrama de las nuevas etapas de entrenamiento y clasificación. La *imagen de entrenamiento* es la que resulta de la separación y mezcla (ver capítulo 9), que contiene muestras de todas las clases. Cada procedimiento de clasificación se caracteriza por:

Figura 14: Diagrama del nuevo esquema de entrenamiento y clasificación.

- el entrenamiento,
- la definición de las clases,
- la asignación de membrecías.

Utilizando los datos de la imagen de entrenamiento y del archivo de configuración, el entrenamiento concluye con la definición de las clases. Luego, cada pixel obtiene una membrecía a cada clase. Finalmente, se asigna el pixel a una sola clase en base al análisis de las membrecías. En la práctica se siguiendo un criterio muy simple: la clase con mayor membrecía.

10.1. Modelo para representar imágenes y sus propiedades

Una imagen se representa con un conjunto de pixels: $I = \{x_1 \dots x_N\}$.

Sean f_1, \dots, f_M las *características*.

Cada pixel se modela mediante el vector de sus *características*:

$$\mathbf{x} = [f_1(x), \dots, f_M(x)]$$

Una clase es un subconjunto de pixels de la imagen. La imagen de entrenamiento define un conjunto $C = \{c_1, \dots, c_n\}$ de clases disjuntas:

$$\forall i = 1 \dots n : c_i = \{x_1^{c_i}, \dots, x_{s_i}^{c_i}\} \subset I$$

\wedge

$$\forall c, s \in C : c \neq s \Rightarrow \forall i, k : x_j^c \neq x_k^s$$

Luego, definimos:

- $\bar{c} = [\bar{f}_1^c, \dots, \bar{f}_{t_c}^c]$ es el denominado *prototipo* de la clase c donde

$$\bar{f}_k^c = \frac{\sum_j f_k^c(x_j^c)}{s_c}$$

- Σ_c es la **matriz de covarianza**, donde

- $q_{ij} = \frac{1}{s_c - 1} \sum_{k=1}^{t_c} (f_i^c(x_k^c) - \bar{f}_i^c)(f_j^c(x_k^c) - \bar{f}_j^c)$

A continuación se describirá cada uno de los métodos por separado.

10.2. Método basado en la distancia Mahalanobis y la distribución χ^2

En las cadenas anteriores, para representar cada clase se utilizan *todas* las *características* y se asume que estas presentan una distribución Normal.

El siguiente método se basa en:

- el hecho que cada clase puede ser representada por un subconjunto particular de *características* y
- se mantienen la asunción de que la distribución de las *características* es Gaussiana en cada clase.

Asumiendo estas cosas, el método se basa en lo siguiente ([n]⁶).

Sean:

$F_c = \{f_1^c, \dots, f_{t_c}^c\}$ el subconjunto de tamaño t_c de *características* que mejor representan la clase c .

$x = [f_1^c(\mathbf{x}), \dots, f_{t_c}^c(\mathbf{x})]$ la representación de x con las *características* de F_c .

Una medida de separación entre el pixel x , y la clase c es la denominada **distancia Mahalanobis** al cuadrado:

$$Z_c^2(x) = (x - \bar{c})^T \Sigma_c^{-1} (x - \bar{c}) \quad (15)$$

Σ_c^{-1} es la inversa de la matriz de covarianza de c o **matriz de precisión** (ver sección 10.1).

Asumiendo que los pixels de una dada clase c tienen una distribución Normal (multivariable), entonces Z_c^2 tiene una distribución χ^2 con t_c grados de libertad (o dof, por sus siglas en inglés *degrees of freedom*).

A partir de esta información el método se implementó de acuerdo con la siguiente sección.

⁶Timothy S. Moore y otros, A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms

10.2.1. Entrenamiento, definición de clases y membrecías

El entrenamiento implica calcular el **prototipo** y la **matriz de precisión** para cada clase. El archivo de configuración provisto debe indicar para cada clase, cuales son las *características* que mejor la representan y diferencian del resto.

Luego, cada clase queda definida a partir de \bar{c} y Σ_c^{-1} .

Finalmente, para un pixel x dado, la membrecía a una clase c se define como:

$$m_c(x) = 1 - F(Z_c^2(x)) \quad (16)$$

donde F es la función de distribución acumulada de una χ^2 con t_c dof.

En la figura 15 pueden verse algunos ejemplos de la forma de esta función.

Figura 15: Ejemplos de la función de distribución acumulada χ^2 con distintos dof.

10.3. Fusión de la información

Los siguientes métodos se basan en el concepto de fusión de la información. Dentro de este marco teórico utilizamos principalmente las llamadas *funciones de masa* (o *masa elemental de creencia*).

Se extenderá el formalismo del capítulo 10.1 basándose en la *Teoría de la Evidencia* de Dempster-Shafer^{7 8}.

⁷Shafer Glenn, A mathematical theory of Evidence, Princeton University Press, 1976

⁸Shafer Glenn, Dempster-Shafer theory, Wikipedia?

Para cada par (clase, *característica*), (c, f) , se define una función de masa, $m_{c,f} : I \rightarrow [0, 1]$. Esta función mide la creencia o *evidencia* que aporta f con respecto a la pertenencia de un pixel a la clase c . La única restricción es que:

$$\forall \mathbf{x}, c, f : 0 \leq m_{c,f}(\mathbf{x}) \leq 1$$

En el marco de la teledetección, la fusión de la información se utiliza principalmente para mezclar los resultados de varios clasificadores y así generar una clasificación final con mejores resultados^{9,10,11}

Para este trabajo se interpreta de una manera algo distinta, tomando cada *característica* como un clasificador (ya que cada *característica* brinda información sobre cada clase, en distinta medida).

En la práctica, para implementar las funciones de masa se utilizó el hecho que las *características* tienen una distribución Normal en cada clase :

Sea $N_{\mu,\sigma}$ la función de densidad de probabilidades Normal con media μ y desvío estándar σ .

Sean $\mu_{c,f}$ y $\sigma_{c,f}$ la media y el desvío estándar de los valores de f en los pixels de la clase c .

Entonces, se define:

$$m_{c,f}(x) = \frac{N_{\mu_{c,f},\sigma_{c,f}}(x)}{N_{\mu_{c,f},\sigma_{c,f}}(\mu_{c,f})} \quad (17)$$

Figura 16: Ejemplo de la función de masa para la clase **Pasturas** y la *característica Amplitud*.

⁹H. Laanaya y otros,

Classification des sédiments marins par fusion de classifieurs binaires SVM, CMM'06 - Caracterisation du milieu marin, 16 - 19 Oct. 2006

¹⁰Pierre Loonis and others, Multi-classifiers Neural Network Fusion versus Dempster-Shafer's orthogonal rule, IEEE International Conference on Neural Networks, 1995. Proceedings, pages. 2162-2165 vol.4, Dec 1995, Australia

¹¹Nikunj C. Oza and Kagan Tumer, Key Real-World Applications of Classifier Ensembles, Information Fusion - Special Issue on Applications of Ensemble Methods, volume 9, number 1, pages 4-20, 2008

Esta definición de $m_{c,f}$ está motivada en:

- generalidad, ya que sirve para todas las clases y *características*.
- $\forall \mathbf{x}, c, f : 0 \leq m_{c,f}(\mathbf{x}) \leq 1$
- Valores cercanos a la media de la clase devolverán valores altos (cercaos a 1) de masa.
- Valores alejados de la media de la clase devolverán valores bajos (cercaos a 0) de masa.

Por ejemplo, en la figura 16 podemos ver el caso del par (**Pasturas, Amplitud**). Si en el perfil de un pixel, el valor de la amplitud es cercano a 0.35 entonces la *característica Amplitud* brindará mucha evidencia a favor de la clase *Pasturas*. Caso contrario, una amplitud extremadamente baja o demasiado alta (mayor a 0.5) brindará poca (o nula) evidencia a favor de la clase *Pasturas*.

Pero la asunción de Normalidad en la distribución de los valores de las *características* no es necesaria. La forma en que una *característica* representa una clase puede modelarse de manera más precisa. Por ejemplo, en la implementación se utilizaron funciones **sigmoideas**:

$$m_{c,f}(x) = \frac{1}{1 + e^{-a(x-b)}} \quad (18)$$

Donde a y b determinan la posición y forma de la curva y dependen de la clase y *característica* particulares.

Figura 17: Ejemplo de la función de masa para la clase **Floresta** y la *característica Media*.

En la figura 16 podemos ver el caso del par (**Floresta, Media**). Si en el perfil de un pixel, la media del NDVI es mayor a 0.7 entonces la *característica Media* brindará mucha evidencia a favor de la clase *Floresta*. Caso contrario, una media baja (menor a 0.5) brindará poca (o nula) evidencia a favor de la clase *Floresta*.

10.3.1. Método basado en la regla de combinación de Dempster

[Agrego el formalismo y desarrollo la teoría de la Evidencia de Dempster-Shafer?]

La regla de combinación de Dempster sirve para combinar piezas independientes de información obtenida a partir de las evidencias que proveen las funciones de masa.

La regla, adaptada a nuestro modelo, es:

Dada una clase c , la fusión de la información provista por las *características* s y t es:

$$m_{c,s} \otimes m_{c,t} = \frac{m_{c,s}m_{c,t}}{|1 - D_{s,t}|} \quad (19)$$

donde

$$D_{s,t} = \sum_{\substack{q,r=1\dots n \\ q \neq r}} m_{q,s}m_{r,t} \quad (20)$$

es una medida del conflicto entre las masas s y t .

Como establece Smets¹² el operador \otimes es asociativo. Luego, las masas de todas las *características* pueden ser fusionadas en un solo valor, con el siguiente procedimiento recursivo:

Sea:

$$G_{c,2} = \frac{m_{c,1}m_{c,2}}{|1-D_{1,2}|} = m_{c,1} \otimes m_{c,2} \quad (21)$$

$$G_{c,k+1} = \frac{G_{c,k}m_{c,k+1}}{|1-D_{k+1}^c|} = m_{c,1} \otimes \dots \otimes m_{c,k+1} \quad (22)$$

donde

$$D_{k+1}^c = \sum_{\substack{r,s=1\dots n \\ r \neq s}} G_{r,k}m_{s,k+1} , n \in [2; M - 1] \quad (23)$$

Entrenamiento, definición de clases y membrecías Para poder computar las funciones de masa, es necesario tener, para cada clase c y *característica* f , la media y desvío estándar de los valores de f en los pixels de la clase c . Estos datos son estimados en el entrenamiento. El archivo de configuración no es utilizado en este caso.

Para un pixel x dado, la membrecía a una clase c es la fusión de las evidencias aportadas por todas las *características*:

$$\begin{aligned} m_c(x) &= G_{c,M-1}(x) \\ &= m_{c,1}(x) \otimes m_{c,2}(x) \otimes \dots \otimes m_{c,M}(x) \end{aligned} \quad (24)$$

¹²Phillipe Smets, **The combination of evidence in the transferable belief model**, IEEE transactions on pattern analysis and machine intelligence, vol. 12, No. 5, pp. 447-458, Mayo de 1990

10.3.2. Método basado en la fusión mediante una Media Ponderada

Finalmente, utilizando el conocimiento de que algunas *características* ayudan a discriminar ciertas clases más que otras, se implementó otro método de fusión de la información: la media ponderada.

El método es extremadamente simple y eficiente. El problema es que requiere de una serie de valores o *pesos* que miden en qué grado contribuyen las *características* para discriminar cada clase.

En definitiva, para cada par (c, f) , se necesita un peso $0 \leq w_{c,f} \leq 1$ que mide en qué grado aporta f para discriminar la clase c del resto.

Entrenamiento, definición de clases y membrecías Este método utilizando las mismas funciones de masa que el anterior (Ver capítulo 10.3) y practicamente realiza el mismo entrenamiento pero además utiliza una *matriz de pesos* que debe ser provista en el archivo de configuración:

$$\begin{matrix} w_{c_1, f_1} \dots w_{c_1, f_M} \\ \dots \\ w_{c_n, f_1} \dots w_{c_n, f_M} \end{matrix}$$

Luego, para un pixel \mathbf{x} dado, la membrecía a una clase c se define como:

$$m_c(\mathbf{x}) = \frac{\sum_{i=1}^M (m_{c, f_i}(\mathbf{x}) * w_{c, f_i})}{N_c} \quad (25)$$

donde

$$N_c = \#\{w_{c, f} | w_{c, f} > 0\} \quad (26)$$

Es la cantidad de *características* con peso mayor a 0.

Como ya se verá en el próximo capítulo, con este último método se pueden obtener excelentes resultados. El problema es que la construcción de la matriz de pesos es un proceso manual que depende del conocimiento de las clases, las características y de la experiencia del usuario. Trabajo futuro podría ser definir criterios para la creación de dicha matriz o diseñar algún procedimiento para generar dichos valores automáticamente, a partir de los datos.

11. Análisis y comparación de resultados

11.1. Matriz de confusión

A continuación se mostrarán ciertos resultados obtenidos con las distintas cadenas.

La forma de analizar los resultados es mediante *matrices de confusión*. Estas se crean a partir de dos conjuntos de clases:

- $C = \{c_1, \dots, c_n\}$, el resultado de una clasificación y
- $V = \{v_1, \dots, v_n\}$, la verdad de campo.

De tal forma que se espera que $\forall k : c_k = v_k$. Esto sucederá en el hipotético caso que la clasificación sea perfecta.

Luego la matriz de confusión será una matriz M de $n \times n$ tal que

$$M_{i,j} = \#(c_i \cap v_j) \tag{27}$$

y se interpreta como “la cantidad de pixels de la clase j que fueron clasificados como i ”. Entonces, los pixels *bien clasificados* en una clase k serán los que pertenezcan a $c_k \cap v_k$ (representados en M por los elementos de la diagonal).

Figura 18: Matriz de confusión de ejemplo.

En la figura 18 podemos ver una matriz de confusión de ejemplo. Ahí se puede ver que hubieron 9139 pixels bien clasificados como *Agricultura*, pero hubieron 860 pixel clasificados como *Agricultura* que pertenecen a *Pasturas*.

11.1.1. Medidas de precisión

A partir de los datos de la matriz de confusión pueden calcularse ciertas medidas de precisión que informan sobre la fiabilidad de la clasificación. Estas medidas se expresan como porcentajes:

Precisión General (Overall accuracy): Es la proporción del total de pixels **bien** clasificados en el total de pixels clasificados:

$$OA = \frac{\Sigma_D}{T} \quad (28)$$

Donde

- $\Sigma_D = \sum_{i=1}^n M_{i,i}$, es la suma de los elementos de la diagonal. Representa la totalidad de pixels bien clasificados.
- $T = \sum_{i=1}^n \sum_{j=1}^n M_{i,j}$ es el total de pixels clasificados.

Es una medida del desempeño general del clasificador.

Precisión del Productor (Producer's accuracy): Se realiza una lectura de la matriz **por columnas**. Para una clase k determinada, la cantidad de pixels bien clasificados como k sobre la totalidad de pixels k en la verdad de campo se denomina *Precisión del Productor*:

$$PA_k = \frac{M_{k,k}}{V_k} \quad (29)$$

Donde

- $V_k = \sum_{i=1}^n M_{i,k}$ es la cantidad de pixels en v_k .

Mide que tan bien se discrimina la clase k o la capacidad del clasificador de identificar los pixels de la clase k .

Precisión del Usuario (User's accuracy): Se realiza una lectura de la matriz **por filas**. Para una clase k determinada, la cantidad de pixels bien clasificados como k sobre la totalidad de pixels clasificados como k se denomina *Precisión del Usuario*:

$$UA_k = \frac{M_{k,k}}{C_k} \quad (30)$$

Donde

- $C_k = \sum_{j=1}^n M_{k,j}$ es la cantidad de pixels en c_k .

Es una medida de qué tan bien representa la clase k la verdad de campo.

Coefficiente Kappa: Algunos pixels pueden estar bien clasificados por una cuestión de *suerte* o de *casualidad*. El coeficiente Kappa intenta brindar una medida general del desempeño del clasificador, teniendo en cuenta ese factor de azar o casualidad.

$$\hat{K} = \frac{T * \Sigma_D - \sum_{k=1}^n (V_k * C_k)}{T^2 - \sum_{k=1}^n (V_k * C_k)} \quad (31)$$

11.2. Resultados

Debe tenerse en cuenta que durante la duración de la pasantía que dio origen a este trabajo, se realizaron muchísimas clasificaciones y pruebas, con distintos conjuntos de *características* y de clases. Aquí solo se presentarán algunos resultados finales, comparables en función de que se realizaron las pruebas con el mismo conjunto de *características* y clases.

Primero se compararán los resultados de una cadena original y una cadena mejorada. Finalmente se analizarán los mejores resultados obtenidos con los nuevos métodos de clasificación con las mismas condiciones que los anteriores.

11.2.1. Condiciones generales de las pruebas

[características, conjunto de clases. Explicar trade-off entre número de clases y performance general: algunas clases son demasiado chicas para un pixel de 250 mts (agua), algunas tienen pocos pixels (Eucaliptus). explicar matriz asimétrica y fusión de clases para analizar los resultados]

11.2.2. Cadena original

11.2.3. Cadena mejorada

11.2.4. Mahalanobis y χ^2

11.2.5. Regla de combinación de Dempster

11.2.6. Promedio ponderado

[Hacer notar la diferencia de performance entre Urbana y Florestas, por ejemplo]

12. Conclusión y trabajo futuro

La deforestación puede buscarse en un paso posterior, con el criterio encontrado.

La clase Urbana puede sacarse de la clasificación y discriminarse mediante máscaras o datos (geográficos) externos.

La matriz de pesos debería poder calcularse usando información de los histogramas de las *características*.