

REPORTE TÉCNICO
**Reconocimiento
de Patrones**

**Estado actual de los métodos
automáticos de anotación semántica
de imágenes de teledetección
espacial (satelitales y aéreas)**

Ing. Alain Jorge Acuña,
Dr. C. Eduardo Garea Llano

RT_033

octubre 2010





CENATAV

Centro de Aplicaciones de
Tecnologías de Avanzada
MINISTERIO DE LA INDUSTRIA BÁSICA

RNPS No. 2142
ISSN 2072-6287
Versión Digital

SERIE AZUL

REPORTE TÉCNICO
**Reconocimiento
de Patrones**

**Estado actual de los métodos
automáticos de anotación semántica
de imágenes de teledetección
espacial (satelitales y aéreas)**

Ing. Alain Jorge Acuña,
Dr. C. Eduardo Garea Llano

RT_033

octubre 2010



Estado actual de los métodos automáticos de anotación semántica de imágenes de teledetección espacial (satelitales y aéreas)

Ing. Alain Jorge Acuña, Dr. C. Eduardo Garea Llano

Centro de Aplicaciones de Tecnología de Avanzada, 7a #21812 e/ 218 y 222, Siboney, Playa, Habana, Cuba.
ajorge@cenatav.co.cu

RT_033 CENATAV

Fecha del camera ready: 30 de junio de 2010

Resumen: La teledetección es una técnica que tiene como objetivo la captura, tratamiento y análisis de imágenes de la superficie terrestre; estas imágenes se encuentran involucradas en una gran variedad de aplicaciones que van desde el estudio y solución de problemas comunes hasta otros sumamente complejos. Con la creciente evolución y puesta en explotación de imágenes satelitales de alta resolución multispectrales e hiperspectrales, grandes volúmenes de archivos de imágenes son capturados y almacenados diariamente y el acceso rápido y efectivo al contenido de estas imágenes a través de las redes y en ambientes distribuidos constituye una tarea engorrosa. La anotación de dichas imágenes es una solución sólida para acceder de manera segura al contenido de las mismas. El término anotación de imágenes ha venido evolucionando desde el simple etiquetado hasta el punto de que ya se ha comenzado a utilizar el término anotación semántica de imágenes, sobre la base del uso de estructuras computacionales que representen formalmente el conocimiento expresado en las mismas, como son las ontologías. En este trabajo se realiza un análisis crítico de los principales métodos automáticos realizados en el campo de la anotación semántica de imágenes y como los mismos están siendo utilizados para anotar semánticamente imágenes de teledetección, así como las perspectivas y problemas que en este sentido se deben ir atacando en las investigaciones venideras.

Palabras clave: auto-anotación semántica, ontologías, imágenes de teledetección espacial

Abstract: Remote sensing is a technique that aims to capture, processing and analysis of images of the earth's surface, these images are involved in a variety of applications ranging from the study and solution of problems common to other highly complex. With the growing trend and start exploiting high-resolution satellite imagery from multispectral and hyperspectral, large volumes of image files are captured and stored daily, and the quick and effective access to the content of these images through networks and distributed environments is cumbersome. The annotation of such images is a solid solution to securely access the content thereof. The term annotation of images has evolved from simple labels to the point that has already begun to use the term semantic annotation of images, based on the use of computational structures that represent knowledge formally expressed therein, such as ontologies. In this paper a critical analysis of the major automated methods performed in the field of semantic image annotation and how they are being used to annotate semantically remote sensing images, and the prospects and problems in this regard should be attacking in future research.

Keyword: Semantic Auto-annotation, Ontology, Images of Space Remote Sensing

1. Introducción

La teledetección es una técnica que tiene como objetivo la captura, tratamiento y análisis de imágenes de la superficie terrestre; estas imágenes se encuentran involucradas en una gran variedad de aplicaciones que van desde el estudio y solución de problemas comunes hasta otros sumamente complejos como son el desarrollo y planificación urbana, la infraestructura, el uso de los suelos, los estudios ambientales y de los recursos naturales, la agricultura, los desastres naturales, la definición de los diferentes tipos de vegetación, la determinación del avance de zonas agrícolas entre otros. El volumen de datos adquiridos mediante la

teledetección crece gradualmente con el avance del tiempo y de las propias técnicas de teledetección, y junto con el incremento de esta cantidad de archivos de imágenes que son obtenidos diariamente, el acceso al contenido de las mismas se dificulta cada vez más y el proceso de anotación semántica de imágenes de forma manual (baja eficiencia e inconsistente el proceso de etiquetado de las imágenes) se ha convertido en un procedimiento engorroso y de difícil realización, a esto se une también la preferencia de los usuarios en cuanto a su inclinación por la realización de consultas basadas en semántica antes que las consultas basadas en contenido.

La mayoría de los actuales sistemas de recuperación de imágenes basados en contenido usan directamente valores simbólicos y proporcionan resultados satisfactorios. Sin embargo, estos rasgos simbólicos no satisfacen completamente las expectativas de los usuarios, debido a que los mismos piensan en términos de conceptos semánticos (“área industrial”, “suburbio residencial”) y no en términos de extracción de valores simbólicos (“textura rayada”, “área verde”) y además todas estas técnicas se basan en la realización de consultas basadas en contenido usando conceptos que muchas veces los usuarios no están familiarizados con ellos. Con el objetivo de incrementar la explotación actual de las imágenes de teledetección es necesario establecer sistemas que posibiliten el acceso selectivo y flexible al contenido de la información de los archivos de imágenes según Bordes et al [1]. La anotación semántica de imágenes despierta gran interés entre los investigadores debido a su amplia utilidad, principalmente en la recuperación de imágenes por interpretación semántica.

2. Definiciones

Anotación: Información textual para describir un objeto o para proporcionar información adicional (ver también etiqueta) [2].

Etiqueta: Nombre o descripción textual de un objeto geográfico [2].

Semántica: El término semántica se refiere a los aspectos del significado, sentido o interpretación del significado de un determinado elemento, símbolo, palabra, expresión o representación formal. En principio cualquier medio de expresión (lenguaje formal o natural) admite una correspondencia entre expresiones de símbolos o palabras y situaciones o conjuntos de cosas que se encuentran en el mundo físico o abstracto que puede ser descrito por dicho medio de expresión [3].

Ontología: El término *Ontología* proviene del campo de la filosofía y la epistemología, donde se define como una rama de la metafísica que se ocupa del estudio de la naturaleza de la existencia, de los seres y de sus propiedades transcendentales. A finales de la década de los 80s el término es adoptado por la Inteligencia Artificial (IA) con el objetivo de compartir y reutilizar conocimiento. Gruber [4] definió ontología como “una especificación explícita de una conceptualización”. Basándose en la definición dada por Gruber, Borst [5] plantea que: “las ontologías son definidas como una especificación formal de una conceptualización compartida”. Guarino [6] creó una distinción entre una ontología y una conceptualización y comenzó el debate diciendo que una conceptualización es “un conjunto de relaciones conceptuales definidas sobre un espacio de dominio” y es importante “centrarse en el significado de estas relaciones, con independencia del contexto”. En Smith et al [7] se plantea que en las investigaciones más recientes sobre intercambio de información, una ontología es vista como un diccionario de términos expresado en una sintaxis canónica.

Anotación semántica: Existen varias definiciones para el término anotación semántica. En Klien et al [8] se define anotación semántica como la correspondencia entre elementos entre un esquema de datos y elementos de una ontología de dominio. En Grcar et al [9] se define como la tarea de hacer coincidir una palabra arbitraria o término con el concepto más apropiado en la ontología de dominio que corresponda. En el proyecto SWING (*Semantic Web Services Interoperability for Geospatial Decision Making* (FP6-026514)) (SWING 2008) se define como el establecimiento de vínculos explícitos entre la información

geográfica que se sirve a través de Web Services OGC y el vocabulario definido y compartido en una ontología de dominio según Grcar et al [9].

Anotación automática de imagen (conocida también como etiquetado automático de imagen): Consiste en el proceso mediante el cual un sistema automatizado asigna metadatos en forma de palabras claves o captioning a una imagen digital. Esta aplicación consiste en una técnica del campo de visión por computadora que es utilizada en los sistemas de recuperación de imagen para organizar y localizar imágenes de interés desde una base de datos [10].

Teledetección: Conocida también como percepción remota, la teledetección es la técnica que permite obtener información sobre un objeto, superficie o fenómeno a través del análisis de los datos adquiridos por un instrumento que no está en contacto con él. Se basa en que cada objeto, área o fenómeno emite un espectro electromagnético específico, en función de su propia naturaleza y de las radiaciones que recibe. La reflectancia de ese espectro electromagnético se denomina firma espectral, la cual hace distinguible a ese objeto, superficie o fenómeno de los demás. Por lo general los datos son recogidos a través de sensores instalados en plataformas aerotransportadas o en satélites artificiales, los cuales captan la radiancia emitida o reflejada, obteniéndose una imagen, habitualmente en falso color con una banda para cada una de estas regiones del espectro. Los avances en tecnología han permitido contar con instrumentos cada vez más precisos basados en electrónica y experimentación con materiales que permiten obtener información cada vez más completa contenida en imágenes satelitales según Gutiérrez et al [11].

3. Técnicas para la anotación automática de imágenes

Para automatizar el proceso de anotación de imagen se procede a la definición de pasos de extracción de rasgos y generación de modelos como se refleja en la figura 1. De esta manera modificando cualquiera de los dos pasos se obtiene un nuevo método de anotación. Por consiguiente, combinando cualquier método de generación de rasgos y técnicas de aprendizaje de máquinas se puede construir un sistema de anotación automática según Zhang [12].

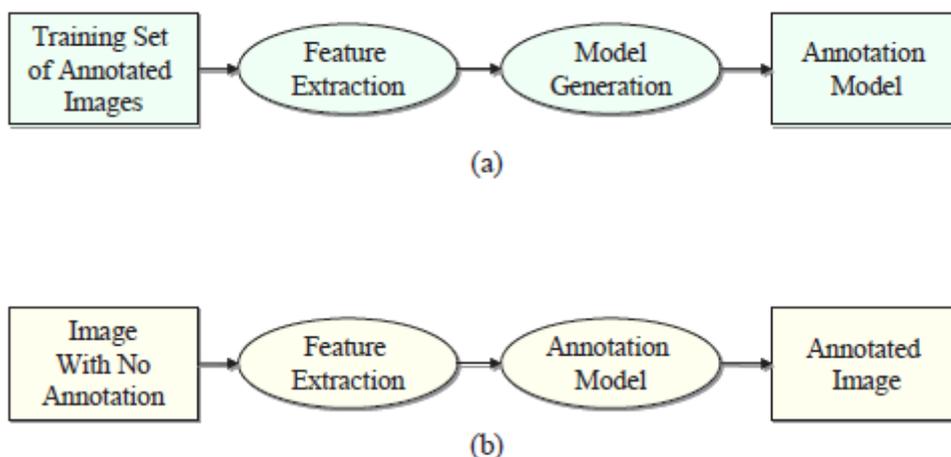


Fig.1. (a) Generando un modelo de anotación; (b) Anotando una imagen según [12]

En la figura 2 se presenta una clasificación de los métodos de anotación de imágenes según Zhang [12]. Debido a la naturaleza ineficiente de los métodos manuales de anotación

de imágenes; en este trabajo nos enfocamos en el estudio de los métodos automáticos o semiautomáticos de anotación de imágenes.

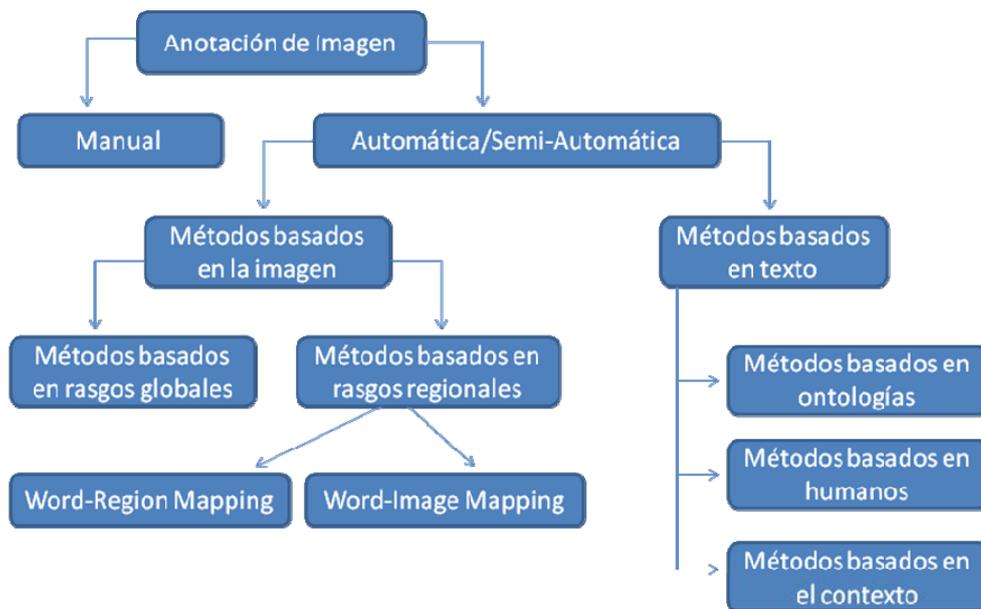


Fig. 2. Clasificación de los métodos de anotación de imágenes según [12]

3.1 Semántica en las imágenes

Según [19] la semántica de una imagen define el significado que tiene ésta. Las formas de los objetos contenidos en una imagen son una característica importante del contenido semántico de la imagen y puede dar como resultado un índice en conjunto con alguna métrica de categorización estableciendo una anotación inteligente para su representación y recuperación. La semántica de una imagen describe sus atributos tales como tipos de objetos, objetos individuales y eventos, acciones, emociones o incluso algún significado religioso.

Estrictamente hablando de la semántica de la imagen apelaría a la explicación de las estructuras y gramáticas que dan significado a los textos visuales. La semántica de las imágenes apela al contexto, a los usos, a lo simbólico, y no solo a las categorías y taxonomías de tipo estructural o lógico según Lizarazo [20].

3.2 Métodos actuales para la extracción de rasgos semánticos en las imágenes

Zhang [12] explica que los métodos actuales para la extracción automática de rasgos semánticos desde imágenes se pueden dividir en dos grupos: los métodos basados en texto que hacen uso de metadatos tales como descripciones ontológicas y/o texto asociado con imágenes para asignar y/o refinar las anotaciones y los métodos basados en la imagen que se centran en la extracción de información semántica directamente de la imagen. Hacen referencia a la importancia de la anotación automática de las imágenes y su uso en diferentes campos como la biomedicina, comercio, periodismo, en el campo militar, en la educación, librerías digitales, búsquedas webs, en la publicidad, etc.

Según Zhang [12] la clasificación de los métodos de extracción de rasgos semánticos tiene dentro de sí misma varias sub-clasificaciones que se describen a continuación:

3.2.1 Métodos basados en la imagen (Image-based methods)

Muchos estudios se centran en la extracción de la información semántica directamente desde el contenido de la imagen. Esto convierte el problema de la anotación en un problema de clasificación semántica de las imágenes. Ahora los métodos de anotación basados en la imagen se pueden clasificar en:

3.2.1.1 Métodos basados en rasgos regionales (Regional feature-based methods)

La mayoría de los métodos de anotación de imagen se basa en la extracción de rasgos regionales de las imágenes, que es, de hecho basarse en segmentos semánticamente significativos o regiones de la imagen. Las técnicas de segmentación y/o partición de una imagen son utilizadas para la detección de objetos dentro de la misma imagen. Luego se generan modelos para calcular la probabilidad de co-ocurrencias de palabras con rasgos regionales de la imagen. Estos modelos son utilizados para anotar nuevas imágenes [21], [22], [23], [24], [25], [26], [27].

3.2.1.2 Métodos basados en rasgos globales (Global feature-based methods)

Estos métodos proceden a la extracción de rasgos globales de la imagen, los cuales tienen como desventaja frente a los rasgos regionales que no capturan información detallada con respecto a objetos individuales. Por esta razón algunos tipos de anotaciones pueden no ser posibles de realizar [23], [24], [28], [29].

3.2.2 Métodos basados en texto (Text-Based Methods)

Estos métodos hacen uso de conocimiento adicional acerca de las imágenes que existe en textos asociados, ontologías para la realización de la anotación. De manera general, este tipo de anotación basada en texto puede dividirse en tres principales categorías:

3.2.2.1 Métodos basados en ontologías (Ontology-based methods)

Estos métodos se apoyan en una ontología que puede estar formada por una lista de palabras claves, categorías jerárquicas o bases de datos lexicográficas. Para generar una anotación se hace uso de la asociación de texto con la imagen usando los términos ontológicos, reglas y/o relaciones presentes en la o las ontologías [30], [31], [32], [33], [34], [35], [36], [37], [38].

3.2.2.2 Métodos basados en el contexto (Context-based methods)

Una imagen puede estar presente en el contexto de un documento multimedia. Por ejemplo una imagen en una página web puede estar rodeada por texto, o un frame de imagen en un clip de video pudiera tener captions de texto y/o audio. Tal contenido contextual de los documentos multimedia pudiera ser utilizado para anotar imágenes [39], [40], [41], [42].

3.2.2.3 Métodos basados en humanos (Human-based methods)

Estos métodos prestan ayuda a los usuarios para asignar anotaciones utilizando métodos que son diseñados con el propósito de disminuir la cantidad de operaciones necesarias para la anotación según Walter et al [43].

A continuación abordaremos con más detalles cada uno de los métodos de anotación de imágenes planteados por Zhang [12].

3.2.3 Métodos basados en ontologías (Ontology-based methods)

Según Zhang [12] una ontología brinda una manera conveniente para caracterizar un dominio determinado mediante la conceptualización y especificación de conocimiento en términos de entidades, atributos y relaciones. Las ontologías nos permiten distinguir entre una cantidad determinada de varios tipos de objetos y describir las relaciones entre los mismos, las dependencias y propiedades. Una ontología construye la estructura relacional de conceptos que uno puede utilizar para describir y razonar acerca de aspectos de la vida cotidiana. “Ellas pueden jugar un papel de representación pasiva o rol taxonómico, en términos de compromisos ontológicos previos para ciertos objetos y categorías y representan un rol activo inferencial que define cómo pueden ser identificadas las propiedades relevantes de nuestro mundo” según Town [44].

El proyecto World Wide Web Consortium (W3C) se refiere a la Web Semántica como un nuevo espacio que brinda a los usuarios la posibilidad de encontrar, compartir y combinar más fácilmente información basándose en la información legible para las máquinas y en metadatos. De esta iniciativa han surgido variedades de ontologías, ya sea de propósito general como estándares de dominios específicos, así como herramientas como Lenguaje Web de Ontologías (OWL) (<http://www.w3.org/TR/owl-features/>) y un framework para la descripción de recursos (<http://www.w3.org/TR/owl-features/>).

Muchos investigadores como Hu et al [45] y Soo et al [46], han demostrado el uso de ontologías especializadas en la medicina, arte e historia con el fin de llevar a cabo anotaciones de dominios específicos.

3.2.4 Métodos basados en el contexto (Context-based methods)

Muchos estudios se han dirigido hacia estos métodos basados en el contexto, especialmente con respecto a la Word Wide Web. Buenos ejemplos basados en el contexto de automatización de anotación de imagen basada en la Web y sistemas de recuperación incluyen los motores de búsquedas tales como Google Image Search (<http://images.google.com>). Estos motores de búsquedas analizan el texto en las páginas web adyacentes a una imagen, sus captions y otros factores con el fin de determinar el contenido de la imagen.

El “weight chainNet model” propuesto por Shen et al [47] es uno de los trabajos representativos para la anotación y recuperación de imágenes en la Web. Está basado en el hecho que una imagen en una página Web está relacionada semánticamente con su texto circundante. Se identifican cuatro partes del contenido textual que están bien embebidas con la imagen, estas partes son: image title, image ALT (alternate text), image caption y page title. Este modelo está basado en el uso de cadenas léxicas (lexican chains) que representan la semántica de una imagen desde su texto más cercano.

3.2.5 Métodos basados en humanos (Human-based methods)

De manera diferente a los métodos antes analizados, estos métodos son utilizados para imágenes que no tengan contenido de texto asociado (el contexto de la imagen se utiliza en un porción mínima o no se utiliza). Usualmente las interfaces de usuarios facilitan y minimizan las operaciones manuales, con el fin de mejorar la exactitud y eficiencia de las anotaciones. En Lieberman et al [48] se presentaron una de estas interfaces de sistemas.

3.2.6 Métodos basados en rasgos regionales (Regional feature-based methods)

Partiendo de un conjunto de entrenamiento de imágenes anotadas, los métodos basados en rasgos regionales intentan descubrir vínculos estadísticos entre rasgos visuales y palabras

seguido por una estimación de la correlación entre palabras y rasgos regionales de una imagen.

Dado un conjunto T de imágenes de entrenamiento, uno o más de los siguientes pasos son utilizados con este fin:

- **Segmentación:** Los algoritmos de segmentación de propósito general como los presentados por Carson et al [49] y Shi et al [50] son utilizados con el propósito de dividir las imágenes en regiones, seguido por la extracción y cuantificación de rasgos desde las regiones. Estas regiones son descritas por un conjunto de rasgos de bajo nivel como color, textura, forma y así sucesivamente. La segmentación de la imagen influye ampliamente en la calidad de la anotación automática de la imagen. Ninguno de los métodos existentes en la literatura garantiza la creación de regiones semánticamente significantes, sin embargo las regiones obtenidas mediante los mismos son suficientes para el proceso de anotación. Alternativamente uno podría particionar una imagen en un conjunto fijo de regiones rectangulares que evitarían el costo computacional del proceso de segmentación.
- **Agrupamiento:** Las regiones obtenidas del proceso de segmentación son agrupadas en un conjunto de regiones similares, cada grupo (cluster) es denominado blob. La hipótesis aquí consiste en que regiones similares pertenecen al mismo cluster. Debido a que la mayoría de los métodos basados en regiones dependen del agrupamiento (clustering), el funcionamiento del proceso de anotación está fuertemente influenciado por la calidad del agrupamiento.
- **Correlación:** El último paso consiste en analizar la correlación entre palabras y regiones con el objetivo de descubrir semántica oculta. Esto por lo general constituye la tarea más difícil, ya que conjuntos de datos de imagen no suelen proporcionar la correspondencia explícita entre los dos (palabras-regiones). Los métodos basados en rasgos regionales pueden ser clasificados como *word-region mapping* y *word-image mapping*.

3.2.6.1 Word-Region Mapping

Este método consiste en el proceso de asignación de palabras a regiones de una imagen en lugar de la imagen completa; sin embargo con excepción de un vocabulario pequeño es muy difícil encontrar este tipo de datos de aprendizaje. Con el incremento de la longitud de los datos se convierte en un proceso muy difícil para asignar palabras a regiones individuales de una imagen. Con el propósito de hacer frente a este problema podemos hacer uso del Multiple Instance Learning (MIL) propuesto por Dietterich et al [51].

MIL está propuesto como una variación del aprendizaje supervisado (supervised learning) para problemas con conocimiento incompleto acerca de etiquetas de ejemplos de entrenamiento. En el aprendizaje supervisado, cada instancia de entrenamiento es asignada con una etiqueta discreta o de valor real. En comparación, en MIL las etiquetas son solo asignadas a bolsas de instancias. En el caso binario, una bolsa es etiquetada como positiva si al menos una instancia en esa bolsa es positiva y una bolsa es etiquetada como negativa si todas las instancias en ella son negativas. El objetivo de MIL es clasificar bolsas de instancias no vistas basándose en el etiquetado de bolsas como datos de aprendizaje presentado por Yang [52].

En el proceso de anotación de imagen basada en rasgos regionales cada región constituye una instancia y el conjunto de regiones desde la misma imagen representa una bolsa.

Alternativamente, uno podría utilizar el algoritmo Expectation Maximization (EM) propuesto por Moon [53], que es un método de optimización iterativa para encontrar las estimaciones de máxima probabilidad de parámetros desconocidos en modelos probabilísticos; los parámetros desconocidos son denominados como variables latentes no observadas (unobserved latent variables).

Existen dos modelos de word-region mapping: el modelo de co-ocurrencia (co-occurrence model) presentado por Mori et al [54] que usa MIL y el modelo de traducción de máquina (machine translation model) propuesto por Duygulu et al [55] basado en el algoritmo EM.

3.2.6.1.1 Co-Occurrence Model

El modelo de co-ocurrencia presentado por Mori et al [54] para automatizar la anotación de imagen colecciona la cantidad de ocurrencia entre palabras y rasgos regionales y los utiliza para anotar nuevas imágenes. La idea es reducir el ruido acumulando patrones parciales similares desde múltiples imágenes anotadas. En este modelo se hace uso del MIL como se muestra en este ejemplo:

Supongamos una imagen $I_1 \in T$ que está anotada con dos palabras –cielo y montaña- es segmentada en dos regiones: R_{11} (contiene cielo) y R_{12} (contiene montaña). De acuerdo con MIL, cada región de una imagen es asignada con palabras que son heredadas desde la imagen completa (así R_{11} y R_{12} son anotadas con cielo y montaña). Note que la palabra montaña representa una anotación inapropiada para R_{11} . Sin embargo, si tenemos otra imagen $I_2 \in T$ que es también anotada con dos palabras –cielo y río- que también es dividida en dos regiones R_{21} (contiene cielo) y R_{22} (contiene río), luego acumulando las palabras desde estas dos imágenes anotadas, tenemos dos patrones cielos, una montaña y un río. De tal manera, MIL asume que la tasa de palabras inapropiadas decrece gradualmente acumulando patrones similares.

La probabilidad $P(w_i | b_j)$ para cada palabra w_i y blob b_j es estimada acumulando su frecuencia de ocurrencia como sigue:

$$\begin{aligned}
 P(w_i | b_j) &= \frac{P(w_i | b_j)P(w_i)}{\sum_{k=1}^n P(b_j | w_k)P(w_k)} \text{ por el teorema de Bayes} \\
 &\approx \frac{\binom{m_{ji}}{n_i} \binom{n_i}{N}}{\sum_{k=1}^n \binom{m_{jk}}{n_k} \binom{n_k}{N}} \\
 &= \frac{m_{ji}}{\sum_{k=1}^n m_{jk}},
 \end{aligned}$$

donde n es el número de palabras del vocabulario, m_{ji} es el número de ocurrencias de la palabra w_i en el blob b_j , n_i es el número de ocurrencias de la palabra w_i en todos los blobs, y N es la suma de los números de ocurrencias de todas las palabras en todos los blobs.

Entrenamos un modelo estadístico usando el vocabulario desde todas las imágenes en el conjunto de entrenamiento, procedemos a anotar una nueva imagen, primeramente clasificando sus segmentos para encontrar los blobs más familiares y entonces encontrar las correspondientes palabras para cada blob, seleccionando las palabras con mayor probabilidad.

En Mori et al [54] se utiliza un amplio dominio de imágenes y de vocabulario de palabras durante el experimento realizado con las imágenes a anotar obteniéndose un porcentaje de palabras coincidentes con las imágenes a anotar no muy deseable, pudiéndose mejorar este resultado restringiendo el dominio utilizado en el experimento, también cabe mencionar la carencia en este método de la presencia de un buen sistema de interfaz de usuario que apoye al mismo en la minería de datos.

3.2.6.1.2 Machine Translation Model

El modelo de traducción de máquina presentado por Duygulu et al [55] constituye una mejora sustancial sobre el modelo de co-ocurrencia. Este modelo trata el proceso de anotación como una tarea de traducción desde un vocabulario de blobs a un vocabulario de palabras, hace uso del algoritmo Expectation-Maximization, que integra el correspondiente aprendizaje en el proceso de anotación de imagen.

El modelo de traducción utiliza un conjunto de entrenamiento para construir una tabla de probabilidad de vínculos de blobs con palabras. Esta tabla representa la probabilidad condicional $P(w_i | b_j)$ de una palabra w_i dado un blob b_j (la probabilidad de traducción de un blob b_j en una palabra w_i). Estas probabilidades son obtenidas por la maximización de la probabilidad $l(T)$ de imágenes anotadas desde el conjunto de entrenamiento T :

$$l(\tau) = \prod_{I \in T} \prod_{i=1}^{q_I} \sum_{j=1}^{P_I} P(w_i | b_j) v b_j$$

donde $v b_j$ es el número de blobs b_j que aparecen en la imagen I . El algoritmo EM es aplicado para encontrar la solución óptima, para esto seguimos la siguiente estrategia iterativa:

- Calcular la expectación (estimación) de la tabla de probabilidad para predecir las correspondencias.
- Utilizar las correspondencias para refinar la estimación de la tabla de probabilidad (ejemplo, encontrar el nuevo máximo).

Una vez obtenida la tabla de probabilidad, el proceso de anotación de imagen es el mismo como en el modelo de co-ocurrencia.

Según Duygulu et al [55] el objetivo de este método consiste en el reconocimiento de objetos como máquinas de traducción (machine translation) y este reconocimiento consiste en un proceso de anotación de regiones de una imagen con palabras. Las imágenes son segmentadas en regiones que las mismas son clasificadas en tipos de regiones utilizando una variedad de rasgos presentes en las imágenes. En este método encontramos un procedimiento de anotación de imagen bastante aceptable que nos facilita enfrentar una variedad de problemas presentes en el tema de reconocimiento de objetos que de otro modo serían inaccesibles de enfrentar, también cabe mencionar que representa un método totalmente agnóstico con relación a los rasgos presentes en las imágenes, ya que es posible hacer uso del mismo para cualquier conjunto de características, incluso para conjuntos de características que varíen con la definición de objeto. Es necesario mencionar la importancia de continuar con la investigación de métodos de agrupamiento de regiones para garantizar que el agrupamiento de regiones sea mejorado por la información presente en las palabras utilizadas en el proceso de anotación, así como que dentro del conjunto de características utilizadas no es posible distinguir entre dos instancias pertenecientes a un mismo concepto (ejemplo, no es capaz de distinguir entre tren y locomotora).

3.2.6.2 Word-Image Mapping

El proceso de asignación de palabras a varias regiones de una imagen puede dar lugar a muchos errores debido a la excesiva segmentación de las imágenes. El proceso de mapeo palabra-image utiliza rasgos regionales para asignar palabras a cada imagen (palabras son asignadas a imágenes completas y no a regiones de una imagen).

Se presentan tres modelos estadísticos de mapeo palabra-imagen (Word-Image Mapping):

1. Modelo de aspectos jerárquicos (hierarchical aspect model) (Barnard et al) [56].
2. Modelo de relevancia (relevance model) (Jeon et al) [57].
3. 2D MHMM (two-dimensional multi-resolution hidden Markov model) (Li et al) [58].

3.2.6.2.1 Hierarchical Aspect Model

El modelo de aspectos jerárquicos según Barnard et al [56] para la automatización del proceso de anotación de imagen está inspirado por el modelo jerárquico generativo Hofmann [59], una combinación jerárquica del modelo de agrupamiento asimétrico que mapea la imagen en clusters y el modelo de agrupamiento simétrico que modela la distribución conjunta de las imágenes y sus características. Con su estructura jerárquica, el modelo está bien conformado para las tareas de recuperación de información tales como exploración de base de datos y búsquedas de imágenes basadas en texto y/o en rasgos de imagen.

Este método presentado en Barnard et al [56] constituye un modelo estadístico que no posibilita organizar colecciones de imágenes integrando la información semántica proporcionada por el texto asociado a las imágenes y la información visual proporcionada por los rasgos característicos de las mismas. Este método es muy prometedor para tareas como la recuperación de información, tales como, la exploración de bases de datos o búsquedas de imágenes en texto y/o rasgos de la imagen. Una dificultad presente en este método se encuentra en el proceso de asignación de palabras a las imágenes, ya que el mismo se basa en el significado semántico del proceso de segmentación de la imagen y se plantea que estos segmentadores utilizados funcionan bien en alguna medida debido a que algunas palabras son satisfactoriamente correlacionadas con segmentos de rasgos de imagen sin embargo, muchas palabras y segmentos solamente adicionan ruido a la imagen a anotar.

3.2.6.2.2 Relevance Model

Los modelos de lenguaje basados en relevancia según Lavrenko et al [60] fueron introducidos para facilitar la expansión de consultas según Zhang [12]. Se hace uso del término relevance model para referirse a un mecanismo que determine la probabilidad $P(w|R)$ de observar una palabra w en documentos relevantes para una información particular necesaria según Lavrenko et al [60].

En este modelo encontramos el Cross-Media Relevance Model (CMRM) que parte de la suposición de tener una colección C de imágenes sin anotar. Cada imagen $I \in C$ está representada por un conjunto discreto de un número de blob $I = \{b_1 \dots b_m\}$. Se asume que existe una colección T de imágenes de entrenamiento anotadas previamente, donde cada imagen $J \in T$ contiene una doble representación en términos de palabras y blobs $J = \{b_1 \dots b_m; w_1 \dots w_n\}$. $\{b_1 \dots b_m\}$ representa los blobs correspondientes a regiones de la imagen y $\{w_1 \dots w_n\}$ representa el conjunto de palabras utilizadas en la anotación de la imagen. El número de blobs y palabras en cada imagen (m y n) puede ser diferente de una imagen a otra imagen. En contraste con el modelo de traducción, no asumimos que existe una correspondencia uno a uno entre los blobs y las palabras en una imagen, solo se asume

que un conjunto de palabras claves $\{w_1 \dots w_n\}$ está relacionado con un conjunto de objetos representados por los blobs $\{b_1 \dots b_m\}$.

El método de anotación de imagen presentado por Jeon et al [57] representa un enfoque para automatizar el proceso de anotación y recuperación de imagen partiendo de un conjunto previo de imágenes de entrenamiento, describiendo posteriormente las regiones presentes en las imágenes como un vocabulario de blobs que son generados desde los rasgos característicos de las imágenes utilizando un método de agrupamiento (clustering). Este enfoque muestra la utilidad del uso de modelos de recuperación de información formal para las tareas de anotación y recuperación de imágenes, sin embargo este método propuesto necesita un mejoramiento con relación al proceso de extracción de los rasgos característicos presentes en las imágenes y también en relación con el uso de rasgos continuos que estén presentes en las mismas con el objetivo de mejorar el proceso de anotación automática de este tipo de ficheros.

Todos los modelos planteados (co-occurrence, machine translation, hierarchical, and CMRM) son discretos y no hacen uso de las características continuas de una imagen. Adicionalmente, se apoyan fuertemente en el agrupamiento (clustering) de vectores de rasgos en blobs. De esta manera, la calidad de estos modelos de anotación es sensible a los errores de agrupamiento y son dependientes a la selección a priori del agrupamiento correcto de granularidad (the right cluster granularity).

También existe el Continuous-Space Relevance Model (CRM) que es un modelo que modela rasgos continuos en las imágenes y no depende del clustering, consecuentemente, CRM no sufre del tema de granularidad según Zhang [12]. Sin embargo es necesario mejorar considerablemente la rapidez de funcionamiento de este modelo de modo que pueda ser probado en amplios conjuntos de datos.

3.2.6.2.3 Two-Dimensional Multi-Resolution Hidden Markov Model (2D MHMM)

Este modelo propuesto por Li et al [61] es diferente de los modelos presentados anteriormente; este enfoque asume que cada imagen tienen asignada una categoría, donde cada categoría corresponde a un concepto y contiene un conjunto de palabras claves. Una palabra clave puede pertenecer a más de una categoría. Las imágenes categorizadas son utilizadas para entrenar un diccionario de modelos estadísticos, con cada modelo representando un concepto. El modelo 2D MHMM representa un tipo de modelo estadístico que es generado para cada categoría; cada modelo determina la probabilidad de que una imagen pertenezca a una categoría basándose en varias texturas regionales y colores extraídos desde la imagen. Además, para cada palabra clave se almacena en un índice global la cantidad de categorías a las que esta palabra clave pertenece. Hasta aquí se concluye con el proceso de generación del modelo.

Para anotar una nueva imagen, el sistema extrae la información del contenido de la imagen en varias resoluciones y utiliza cada modelo para determinar la probabilidad de que la imagen pertenezca a la categoría del modelo. El sistema propuesto y la metodología utilizada presentan varias limitaciones como se reflejan a continuación: las imágenes pertenecientes al conjunto de entrenamiento solamente eran imágenes 2D sin tener presente el tamaño del objeto lo que limita potencialmente la exactitud de los conceptos de aprendizaje pertenecientes al vocabulario utilizado por este método, los resultados obtenidos en los experimentos realizados deben ser interpretados cautelosamente debido a que los propios autores reconocen que el conjunto de imágenes utilizadas para entrenar este sistema no fueron las más ideales además de que para conceptos demasiado complejos el sistema necesita alrededor de más de 40 imágenes de entrenamiento para poder construir un modelo confiable que pueda ser utilizado para el proceso de clasificación, o sea, que mientras más complejos sean los conceptos utilizados en el diccionario de conceptos mayor imágenes se

necesitaran como parte del proceso de entrenamiento y mayor tiempo de ejecución para la CPU que se esté utilizando.

En la figura 3 se presenta un análisis comparativo de los métodos de anotación de imagen utilizando los rasgos regionales de las mismas propuesto por Zhang [12].

	<i>Co-occurrence Model</i>	<i>Machine Translation Model</i>	<i>Hierarchical Aspect Model</i>	<i>Relevance Model</i>	<i>2D MHMM</i>
<i>Mapping</i>	word-region	word-region	word-image	word-image	word-image
<i>Learning Method</i>	MIL	EM	EM	language modeling and EM	markov chain
<i>Specialty</i>	image and region retrieval	image and region retrieval	image browsing and searching	ranked image retrieval	image classification
<i>Scalability</i>	scalable	scalable	scalable up to a few thousand images	scalable	scalable and incremental
<i>Overall Performance</i>	good	better	good	best	better

Fig.3. Análisis comparativo de los métodos basados en rasgos regionales según [12]

3.2.7 Métodos basados en rasgos globales (Global Feature-Based Methods)

Estos métodos proceden de manera diferente a los métodos basados en rasgos regionales, los métodos basados en rasgos globales asumen que el proceso de anotación correcta de una imagen requiere la consideración de la imagen completa en lugar de distintas secciones de la imagen. De esta manera el proceso de anotación utiliza una representación virtual de la imagen de la que se deriva un conjunto de características globales.

Zhang [12] menciona que los resultados actuales publicados relacionados con el proceso de anotación de imágenes utilizan diferentes conjuntos de datos, modelos y rasgos y que debido a esto el proceso de comparación del funcionamiento de los resultados publicados en términos de exactitud y eficiencia se hace bastante difícil, se explica también que para facilitar los estudios de comparación entre los sistemas de anotación de imágenes es necesario que los conjuntos de datos anotados sean desarrollados y puestos a disposición de todos los investigadores de manera que se puedan ejecutar comparaciones que determinen el impacto real de los nuevos aportes sobre los métodos automáticos de anotación en cuanto a la elección de los rasgos característicos, de los modelos y la calidad de los datos. Según Zhang [12] la necesidad de la existencia de conjuntos de datos comunes para avanzar en la efectividad de nuevos métodos automáticos de anotación de imágenes es necesaria por tres razones principales. Primeramente, el rendimiento de un método de anotación es dependiente de las anotaciones realizadas con las imágenes utilizadas como conjunto de datos, es decir, si utilizamos conjuntos de imágenes y anotaciones diferentes podríamos obtener conclusiones contradictorias con el rendimiento de dicho método. Segundo, el costo de la obtención de datos anotados podría ser prohibitivo, lo que representa una barrera para investigadores que se encuentran inmersos dentro de este campo, dando como resultado el retraso de la introducción de nuevas perspectivas. Y por último, la existencia de conjuntos de datos comunes facilita el desarrollo de medidas estandarizadas de la calidad de los datos y el

rendimiento de la anotación, consecuentemente obtendríamos una mejor comprensión del proceso de anotación.

3.3 Técnicas de aprendizaje de máquinas (machine learning) aplicadas al proceso de anotación automática de imágenes

Según Koronacki et al [62] los métodos de machine learning son aplicados con efectividad en muchas ramas de las ciencias de la computación; una de estas ramas en la que encontrados la aplicabilidad de estos métodos es el análisis de imágenes y más específicamente la anotación automática de imágenes. Los métodos de auto-anotación de imagen hacen uso de varios enfoques de machine learning, siendo los métodos de clustering y clasificación los más frecuentemente encontrados en este proceso de anotación de imágenes.

Como bien se refleja en Koronacki et al [62] los métodos desarrollados dentro de machine learning, principalmente los métodos de clasificación y clustering constituyen componentes claves dentro de los algoritmos utilizados en la automatización de la anotación de una imagen, los mismos se mencionan a continuación como sigue: Nonparametric models, Neural networks, Support vector machines, decision trees, Ensembles of classifiers, K-mean clustering, Hierarchical clustering, Dimensionality reduction, Feature space discretization, entre otros.

En Tsai et al [63] se presenta un estudio de 50 sistemas de anotación de imagen publicados desde el año 1997 hasta el año 2006 y de estos se hace un pequeño análisis en cuanto a su procesamiento de imagen y módulos de aprendizaje, el número de palabras claves (keywords) asignadas por imagen así como su escalabilidad (número de palabras claves que el sistema aprende por cada anotación que realiza) y los conjuntos de datos utilizados. También presentan un diagrama de una arquitectura general de un sistema de anotación de imagen que se puede observar en la figura 4.

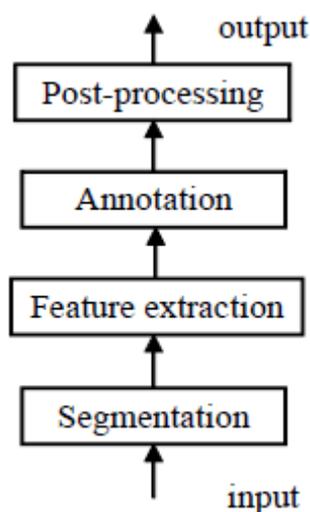


Fig.4. Arquitectura general de un sistema de anotación de imagen según [62]

En Koronacki et al [62] se presentan la mayoría de los métodos de aprendizaje supervisado que apoyan el proceso de anotación de una imagen pertenecientes al conjunto de técnicas de machine learning. A continuación se describen los mismos:

3.3.1 Modelos de aprendizaje supervisado

En Koronacki et al [62] se explica que el aprendizaje supervisado puede ser descrito como un aprendizaje realizado por ejemplos o como un aprendizaje guiado por un profesor y este

profesor tiene el conocimiento del entorno o medio que está representando por un conjunto de ejemplos de entrada-salida. Ahora, con el objetivo de clasificar patrones desconocidos es necesario disponer de un conjunto con cierta cantidad de ejemplos de entrenamiento para cada clase que deseamos identificar y estos datos de entrenamiento son utilizados para entrenar a nuestro clasificador. Luego que tenemos entrenado el clasificador (que sería lo mismo que tener generado nuestro modelo de clasificación) el sistema es capaz de clasificar una instancia desconocida en una de las clases etiquetadas en el conjunto de entrenamiento; más específicamente, el clasificador calcula la similitud entre todas las clases entrenadas y la clase a clasificar y asigna a ésta última una instancia de una clase del conjunto de clases entrenadas con la cual el clasificador identifique que tiene mayor medida de similitud. De este modo el proceso de anotación de una imagen puede ser enfocado como un clasificador que sirva de puente entre el hueco existente entre los vectores de rasgos de bajo nivel y los conceptos de alto nivel (de este modo podríamos corresponder el conjunto de vectores de rasgos de bajo nivel a las clases conceptuales de alto nivel).

3.3.1.1 Clasificadores probabilísticos (Probabilistic Classifiers)

En Koronacki et al [62] se explica que el clasificador Naive Bayes es utilizado para representar la dependencia entre variables aleatorias y para brindar una especificación precisa de la distribución de probabilidad conjunta de un determinado dominio. Este clasificador es construido utilizando datos de entrenamiento para estimar la probabilidad de cada clase dado los vectores de rasgos de una nueva instancia. Por ejemplo, si tenemos el vector de rasgos X el teorema de Bayes nos proporciona un método para calcular la probabilidad que X pertenezca a la clase C_i , representado como $p(C_i | X)$:

$$P(C_i | X) = \prod_{j=1}^N P(x_j | C_i)$$

Esto significa que este clasificador aprende la probabilidad condicional de cada atributo x_j ($j = 1, 2, \dots, N$) de X dada la etiqueta de clase C_i ; por consiguiente el problema de clasificación de imágenes puede ser descrito como un conjunto de rasgos x_j de una imagen X , clasificando X en una de las clases C_i según Koronacki et al [62].

También en este grupo de clasificadores encontramos trabajos que presentan enfoques como esquemas de clasificación jerárquica, modelos ocultos de Markov, modelos ocultos de Markov en dos dimensiones y el modelo LDA (Latent Dirichlet Allocation).

3.3.1.2 Redes neuronales artificiales (Artificial Neural Networks)

Según Koronacki et al [62] las Redes Neuronales (o Redes Neuronales Artificiales) aprenden por experiencia, generalizan el conocimiento de experiencias anteriores a nuevas experiencias y pueden tomar decisiones. Las Redes Neuronales son más flexibles que los clasificadores de tipo Naive Bayes, ya que no necesitamos hacer hipótesis acerca de la distribución de las densidades.

Una Red de Perceptron Multicapa (Multilayer Perceptron Network-MPL) está compuesta por una capa de entrada incluyendo un conjunto de nodos sensoriales como nodos entradas, una o más capas ocultas para los nodos de cómputo y una capa de salida para nodos de cómputo. Los nodos/neuronas de entrada representan los valores de rasgos de una instancia y los nodos/neuronas de salida representan un discriminador entre sus clases y todas las otras clases. Esto significa que cada valor de salida representa una medida de la confianza de red que la clase correspondiente al valor de salida más alto es devuelta como la predicción para una instancia. Cada interconexión tiene asociada consigo un peso escalar que se ajusta durante la fase de entrenamiento.

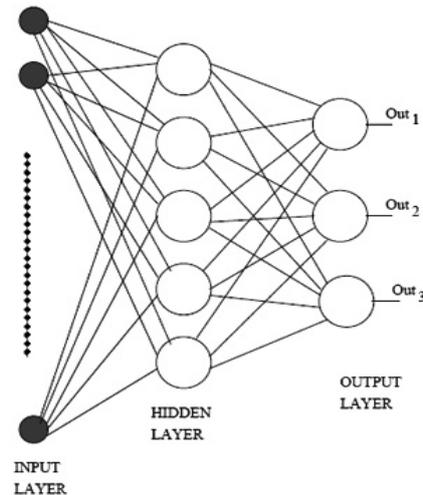


Fig.5. Ejemplo de una red neuronal de tres capas según [62]

En el caso del proceso de anotación de una imagen, los vectores de rasgos de bajo nivel se alimentan dentro de la capa de entrada de un MLP (Multilayer Perceptron Network) donde cada una de las neuronas de entrada corresponde a cada uno de los vectores de rasgos y las neuronas de salida presentes en un MLP representan las etiquetas de clases de las imágenes a ser clasificadas.

3.3.1.3 Máquina de soporte vectorial (Support Vector Machines)

Los Support Vector Machines son diseñados para el proceso de clasificación binaria según Tsai et al [64], durante la generación del modelo SVM (Support Vector Machines), los vectores de entrada (vectores de rasgos de bajo nivel, tales como color y/o textura) son asignados a un nuevo espacio de rasgos dimensional superior denotado como $\Phi : R^d \rightarrow H^f$ donde $d < f$, entonces en el nuevo espacio de rasgos se construye un hiperplano separador. Todos los vectores de un lado del hiperplano son etiquetados como -1 y los que se encuentran del otro lado del hiperplano son etiquetados como +1. Las instancias de entrenamiento que se encuentran más cerca del hiperplano en el espacio transformado son denominadas vectores de apoyo (support vectors) y el número de estos vectores de apoyo es normalmente pequeño comparado con el tamaño del conjunto de entrenamiento y determinan el margen del hiperplano y de este modo la superficie de decisión.

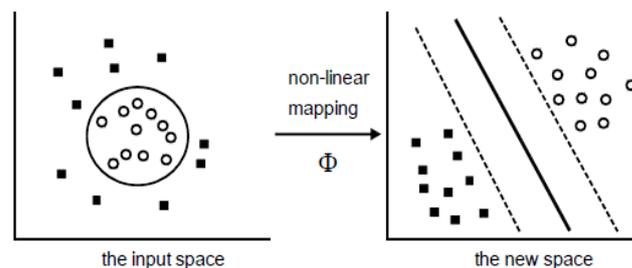


Fig.6. Ejemplo de generación de un modelo de SVM según [64]

3.3.1.4 Árboles de decisión (Decision Trees)

Según Koronacki et al [62] los árboles de decisión son escogidos como clasificadores dentro del proceso de automatización de anotación de imágenes por muchas razones que a continuación se muestran:

- Ofrecen una fase de procesamiento muy rápida con complejidad computacional logarítmica.
- La fase de entrenamiento es también relativamente rápida con respecto a otros enfoques de métodos de aprendizaje de máquinas (machine learning).
- Los algoritmos de entrenamiento no requieren aleatorización (randomization), que no es factible en el caso de la automatización de la anotación de imágenes.
- Los árboles de decisión pueden ser fácilmente reformulados en un conjunto de reglas de decisión.

Mitchell [65] explica que cada nodo de un árbol de decisión representa algún atributo de una instancia mientras que cada rama corresponde a uno de los posibles valores para ese atributo.

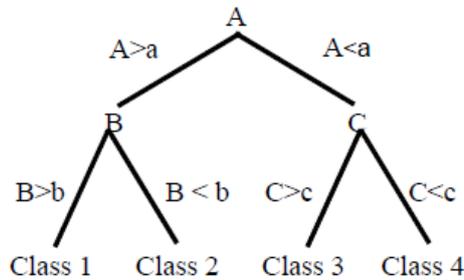


Fig.7. Ejemplo de un árbol de decisión según [63]

3.3.1.5 Vecino más cercano (K-Nearest Neighbor)

Según Bishop [66] el clasificador K-Nearest Neighbor (k-NN) constituye un clasificador no paramétrico convencional. Para clasificar una imagen desconocida representada por algunos vectores de rasgos como un punto en el espacio de rasgos, k-NN calcula las distancias entre el punto y los puntos en el conjunto de datos de entrenamiento, entonces asigna el punto a la clase entre sus k vecinos más cercanos (donde k representa un valor entero) según Tsai et al [63].

3.3.1.6 Template Matching

Este clasificador constituye uno de los primeros y más simples enfoques para el reconocimiento de patrones; el mismo determina la similitud entre dos entidades (tales como puntos, curvas o formas) de un mismo tipo. El patrón/rasgo a ser clasificado es macheado contra una plantilla almacenada (típicamente una forma 2D) y esta plantilla (template) es entrenada partiendo de un conjunto de entrenamiento según Tsai et al [63].

3.3.2 Técnicas de aprendizaje avanzado

3.3.2.1 Combinación de múltiples clasificadores

Según Ghosh [67] y Frosyniotis et al [68] la combinación de un número de clasificadores se ha convertido en un área activa de investigación. Esta combinación suele denominarse conjunto de clasificadores (ensemble classifiers o modular classifiers) y están propuestos con el objetivo de mejorar el funcionamiento de un solo clasificador según Ho et al [69] y Duin et al [70], esto significa que esta combinación es capaz de complementar los errores cometidos por clasificadores individuales en diferentes partes del espacio de entrada.

3.3.2.1.1 Votación mayoritaria (Majority Voting)

El método más simple para combinar clasificadores es votación por mayoría (majority voting), las salidas binarias de k clasificadores individuales se agrupan juntas, luego es seleccionada la clase que recibe el mayor número de votos como la decisión final de clasificación. En general se selecciona la decisión final de clasificación que alcanza la

mayoría de $\frac{k+1}{2}$ de votos.

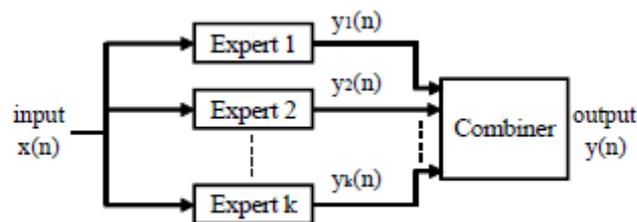


Fig.8. Arquitectura general de un conjunto de clasificadores según Haykin [71]

3.3.2.1.2 Bagging

En esta técnica varias redes son entrenadas independientemente mediante diferentes conjuntos de entrenamiento utilizando el método bootstrap propuesto en Breiman [72]. Bootstrapping construye k replicas de un conjunto de datos de entrenamiento para construir k redes independientes mediante re-muestreo aleatorio pero con sustituciones.

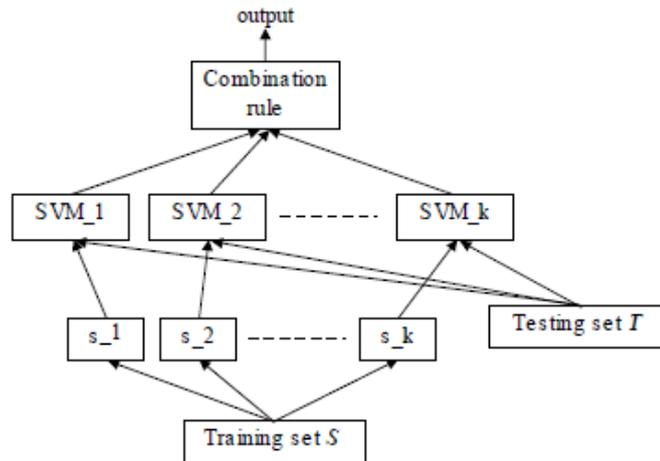


Fig.9. Arquitectura de un módulo de SVM según [63]

3.3.2.1.3 Boosting

En esta técnica como en la anterior, cada red es entrenada utilizando un conjunto de entrenamiento diferente, sin embargo son entrenadas de manera secuencial, no de modo paralelo o independiente según Tsai et al [63]. Schapire [73] propuso el enfoque original de esta técnica, llamado boosting by filtering e intentaba aumentar la exactitud de cualquier clasificador dado (o algoritmo de aprendizaje). Encontramos también el AdaBoost que representa una combinación de ideas detrás de los enfoques de Bagging y Boosting y no necesita de un considerable conjunto de datos de entrenamiento.

3.3.2.1.4 Stacked Generalization

Wolpert [74] propuso esta técnica denominada Stacked Generalization (o stacking) y constituye una vía para combinar múltiples modelos que pueden ser entrenados para funciones de clasificación.

3.3.2.2 Modelos de aprendizaje híbrido (Hybrid Learning Models)

Tsai et al [63] explica que un modelo de aprendizaje híbrido (Hybrid Learning Models) puede estar basado en la combinación de técnicas de agrupamiento (clustering) y clasificación. El agrupamiento puede ser utilizado como un estado de preprocesamiento para identificar clases de patrones para la clasificación supervisada subsecuente según Jain et al [75], que este agrupamiento sería utilizado para una pre-clasificación de colecciones de datos no etiquetadas. Durante el proceso de agrupamiento, estos datos no etiquetados son agrupados y el centro de cada agrupamiento (cluster) puede ser utilizado como ejemplos de entrenamiento para entrenar el modelo de clasificación.

En Tsai et al [63], después del estudio realizado sobre la muestra de 50 sistemas de anotación de imagen publicados desde el año 1997 hasta el 2006 se presenta una comparación de estos sistemas, basándose la misma en el tipo de procesamiento (bajo nivel) de imagen y los modelos de aprendizaje supervisado utilizados por estos sistemas de anotación. Esta comparación no toma en cuenta el funcionamiento del sistema (no se mide ni el proceso de clasificación ni la exactitud de la anotación) debido a que estos sistemas de anotación de imagen utilizan diferentes rasgos de las imágenes, diferentes clasificadores y/o conjuntos de datos y la cantidad de datos de entrenamiento no es la misma por lo que los datos de prueba utilizados reflejan resultados diferentes.

Los sistemas analizados presentan las siguientes características como bien se reflejan en Tsai et al [63] y que se muestran a continuación:

- 21 de los sistemas analizados para indexar imágenes hacen uso de rasgos de las mismas basados en block y cada imagen es clasificada en una sola categoría en específico, 15 sistemas utilizan el procedimiento de segmentación basado en objetos. Encontraron 34 sistemas los cuales asignan una sola palabra clave a cada imagen y solamente 8 sistemas asignan múltiples palabras claves a una imagen utilizando las características locales y regiones de las mismas.
- 14 sistemas hacen uso solamente de rasgos de color para el proceso de aprendizaje de conceptos y para el proceso de clasificación en el marco del problema de escala entre 2 y 50 categorías y solo 6 sistemas tienen en cuenta otros rasgos tales como la forma además de rasgos de color y textura.
- La técnica de Support Vector Machines (SVMs) y los clasificadores modulares se han convertido en un tema de mucho interés recientemente, que se ve reflejado en los 18 sistemas encontrados que utilizan los métodos de aprendizaje supervisado de SVMs y los clasificadores modulares, los métodos de k-NN y Naive Bayes son empleados por 7 y 8 sistemas respectivamente. Los métodos de decision trees y template matching están presentes por separados en un solo sistema de anotación respectivamente.

La gran mayoría de los clasificadores supervisados están implementados para enfrentar el problema de la escala del proceso de anotación a una escala pequeña, por lo que no están diseñados para hacer frente a un proceso de anotación de imagen a gran escala. Es decir, estos métodos están diseñados para clasificar un número pequeño de categorías, todavía constituye un obstáculo la implementación de un clasificador a gran escala. Cuando tenemos en cuenta un gran número de categorías debemos tener presente que los conceptos abstractos y concretos deben ser entrenados para el proceso de anotación de imagen para la recuperación de imágenes por conceptos según Tsai et al [63].

En Tsai et al [63] se menciona que se han identificado tres niveles presentes en un proceso de recuperación de imagen identificados por Eakins [76] y Eakins et al [77], el nivel 1 está basado en rasgos primitivos de las imágenes, tales como color, textura y forma pertenecientes al grado más bajo de abstracción; el nivel 2 está basado en la recuperación de objetos identificados dentro de una imagen e involucra algún grado de inferencia lógica y el nivel 3 está basado en atributos abstractos que involucra un alto grado de razonamiento acerca del significado y propósito de las escenas representadas en las imágenes, tales como eventos nombrados o tipos de actividad, significado simbólico o artístico. Mencionar también que en Tsai et al [63] se explica que la mayoría de los usuarios desean recuperar las imágenes localizando dentro de las mismas, tipos particulares o instancias individuales de un objeto, fenómenos, eventos, así como también reflejan que la mayoría de los sistemas existentes de anotación de imagen involucrados en este estudio realizado desde 1997 a 2006 solamente posibilitan recuperación de imágenes al nivel 2.

A esto se adiciona que la mayoría de las categorías de conceptos necesitan ser clasificadas y que la mayoría de estas categorías existentes son ambiguas (*building/home, men/women/children, rural/park/garden, etc.*) por lo que constituye un tema más desafiante. Existen muchas categorías que son conceptualmente diferentes pero visualmente similares en el espacio de rasgos que pueden provocar solapamiento de rasgos (feature overlapping), más específicamente se podemos categorizar las imágenes en dos tipos de categorías, concreta y abstracta, donde concreta significa un objeto físico o entidad y abstracta significa abstracción, actividad humana o una colección de múltiples objetos/entidades físicas, por consiguiente es necesario entender el funcionamiento de ambos tipos de categorías en procedimientos de anotaciones de imágenes a gran escala según Tsai et al [63].

Otro tema interesante presentado en Tsai et al [63] consiste en la evaluación de los sistemas de anotación de imágenes estudiados en el rango de 1997-2006. Se plantea que para evaluar estos sistemas se procede a seleccionar un conjunto de datos de referencia aplicados como ground truth (), las imágenes de prueba son anotadas de forma manual, donde cada

imagen es asociada a una categoría en específico, luego los sistemas de anotación de imágenes proceden a comprobar si las palabras claves asignadas a las imágenes de prueba coinciden exactamente con sus categorías originales. Ahora, estas palabras claves utilizadas por estos sistemas son inexactas en cuanto a la descripción del contenido de las imágenes y su similitud con relación al mundo real de los usuarios, específicamente porque cada imagen es solamente clasificada en una categoría en específico. En este punto, cabe mencionar que los sistemas de recuperación de imágenes son elaborados para usuarios reales por lo que es necesario la evaluación de los sistemas de anotación de imágenes por parte de los mismos usuarios con el objetivo de entender el funcionamiento de estos sistemas de anotación y no evaluar estos sistemas mediante conjuntos de datos tomados como referencia para a un determinado ground truth.

En Llorente et al [78] se presenta otro estudio acerca de otros enfoques de aprendizaje de máquinas y se reflejan las principales limitaciones de los mismos. Según Llorente et al [78] el principal desafío del proceso de anotación automática de imágenes es la creación de un modelo que sea capaz de asignar términos visuales a una imagen con la intención de describir la misma, se menciona que la mayoría de los algoritmos presentes en la anotación automática de imágenes tienen como punto de partida un conjunto de entrenamiento de imágenes que han sido anotadas previamente por anotadores humanos y que estas anotaciones constituyen metadatos textuales sin estructura constituidos sencillamente de palabras claves (keywords) que describen el contenido de las imágenes. A continuación se muestran los enfoques analizados:

- Co-ocurrence Model
- Machine Translation Model
- Hierarchical Aspect Model
- Cross-Media Relevance Model
- Continuos-Space Relevance Model
- Multiple-Bernoulli Relevance Model
- Latent Dirichlet Allocation
- Probabilistic Latent Semantic Analysis
- Dual Cross Media Relevance Model

Entre las limitaciones encontradas en el estudio realizado en Llorente et al [78] de los enfoques analizados se reflejan que los mismos presentan dos tipos de inexactitudes.

El primer tipo de inexactitud presente en estos métodos corresponde con el problema de reconocimiento de objetos en las escenas presentes en una imagen, se menciona que esto está estrechamente relacionado al uso de la correlación entre rasgos de bajo nivel presentes en las imágenes y las palabras claves (keywords), así también como la dificultad de distinguir entre conceptos que sean visualmente similares, y se refleja que una manera de hacer frente a estos problemas consistiría un refinamiento de los parámetros de análisis de las imágenes. Según Llorente et al [78] en Duygulu et al [79] se pretende hacer frente a este primer tipo de inexactitud; Duygulu et al [79] plantea que estos problemas surgen como resultado del uso de vocabularios no adecuados para fines de investigación y presenta una solución en la que es necesario investigar más en el proceso de traducción que realizan en temas como métodos de agrupamiento de regiones que garanticen que las regiones agrupadas tengan mas relación con las palabras asignadas, así como hacer frente al uso de palabras compuestas dentro del léxico que utilicen para la traducción.

El segundo grupo de inexactitud esta relacionada con la incoherencia entre las etiquetas asignadas como palabras claves a la imágenes anotadas, debido a que las palabras claves son asignadas sin considerar el contexto de las mismas.

Partiendo del estudio realizado en la literatura consultada, mostramos en la figura 10 una clasificación de los métodos automáticos de anotación semántica de imágenes.

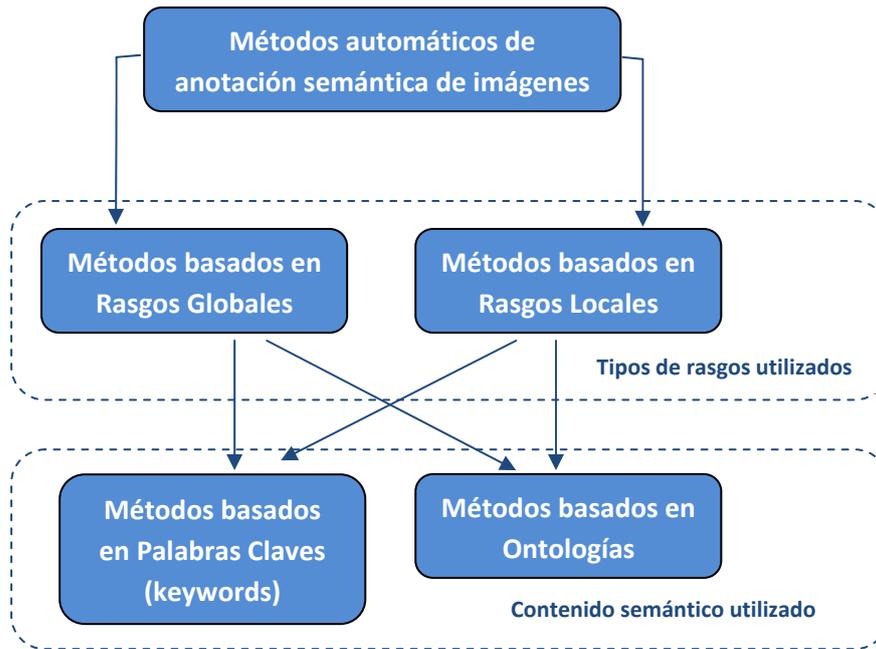


Fig.10. Taxonomía de los métodos automáticos de anotación semántica de imágenes

4. Técnicas para el procesamiento digital de imágenes satelitales

Existen herramientas que brindan a los usuarios la posibilidad de trabajar con la extracción e interpretación de los rasgos y objetos presentes en las imágenes, de clasificar, agrupar y segmentar tipos de escenas presentes en las propias imágenes, de extraer segmentos lineales, detección de bordes, así como la disminución de ruido presente en algunos ficheros de imágenes entre otras muchas operaciones. A continuación mencionamos algunas de estas herramientas.

4.1 Segmentación

El término segmentación suele aplicarse a la búsqueda de áreas homogéneas en una imagen y posterior identificación de las mismas como una clase de información determinada según Mather [80]. La segmentación de imágenes implica la división del territorio en áreas cuyas características espectrales son similares según Fernández et al [81] respecto a ciertas características o atributos.

Es un proceso digital de corte de vértices que consiste en el reconocimiento automático de objetos en una escena sin la intervención del usuario. En el contexto aeroespacial este proceso posibilita el análisis de imágenes aéreas o espaciales, tales como clasificación de diferentes tipos de vegetación para el estudio de su nivel de degradación, clasificación y control de los cultivos de la agricultura, detección de objetos.

En la figura 11 se muestra un ejemplo de un proceso de segmentación aplicado a una imagen de teledetección.

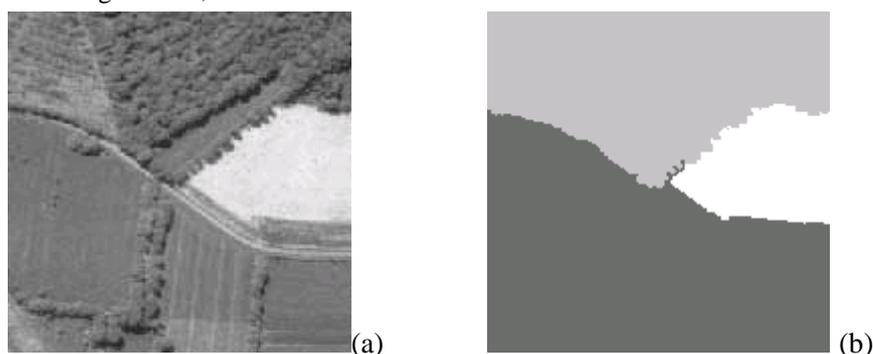


Fig.11. Segmentación de imagen aeroespacial, (a) imagen original, (b) imagen segmentada según Viegas et al [82]

La aplicación de técnicas de segmentación permite extraer características de los objetos detectados y el tipo de características extraídas depende de las técnicas de segmentación aplicadas, las que pueden ser: Umbralización, Detección de bordes, Detección de regiones, Clustering, Segmentación de texturas, Morfología matemática, Redes neuronales, etc. Por un lado se pueden obtener características geométricas de los objetos de la imagen (forma y tamaño) y su localización (altura, ancho, área, perímetro, factor de forma, etc.). Por otro lado, características de intensidad y brillo de la región y su vecindario (valor medio de gris, gradiente medio, características de texturas, contraste, etc.) según Viegas et al [82].

4.2 Detección de objetos

Algunos trabajos se han centrado en la detección de objetos en imágenes satelitales de resolución de 2.5 metros por pixel. Estos son ampliamente aplicables en procesos de clasificación de escenas, clasificación de cromosomas o identificación de objetos. La mayoría de estos métodos definen un conjunto de rasgos para representar la imagen mientras reducen su dimensionalidad según Walker et al [83]. Después que estos rasgos han sido calculados, son usados con una regla de clasificación para establecer una etiqueta a la imagen según Bordes et al [84].

Muchos métodos de detección de objetos se basan en un proceso de aprendizaje que usa un descriptor local que extrae vectores de rasgos en puntos dados de una imagen. Estos vectores de rasgos son entonces comparados con vectores de referencia con el objetivo de detectar los objetos según Bordes et al [84].

En Bordes et al [84] se presenta un estado del arte de los métodos genéricos de detección de objetos. A continuación se presenta una breve descripción del mismo.

- ✓ **Object recognition with rotation invariant moments:** los rasgos algebraicos invariantes de la imagen se definen como radios o momentos de poder de una función de la imagen. Utilización de los momentos de Zernicke y Hilbert, en el caso de los momentos de Zernicke presentan unas cualidades excelentes para describir la forma de los patrones que contienen una imagen: invarianza a la rotación, robustez al ruido, eficiencia expresiva, rápida computación y una representación multi-nivel y en el caso de los momentos de Hilbert.
- ✓ **Object recognition with points of interest:** estos métodos se basan en la extracción de puntos de interés y se calculan los rasgos invariantes a la rotación en estos puntos. Para proceder a la detección de objetos, los puntos en la imagen de aprendizaje son mapeados con los puntos de una imagen de prueba. Una ventaja que presenta este método es que es bastante fuerte frente a modificaciones del background y ocultaciones. Sin embargo, la información que se extrae de los objetos es muy local, debido a que los rasgos solo son calculados en un área muy pequeña cerca de los puntos de interés.

- ✓ **Graph of attributes:** en este enfoque la imagen es primeramente reducida a un conjunto de formas primitivas (segmentos y arcos por ejemplo), luego se procede a la construcción de un grafo cuyos nodos son las formas primitivas y los arcos representan las relaciones espaciales entre ellos (ángulo de intersección, distancia etc.), partiendo de imágenes de aprendizaje se construye el grafo correspondiente que será usado para la detección de objetos en las imágenes. Con este método se extrae un gran cantidad de información de los objetos, pero este tipo de sistema necesita algoritmos complicados de macheo de grafos que dificultan el proceso de establecimiento de tales métodos.

4.3 Clasificación de imágenes satelitales

Gabriel et al [85] define la clasificación de una imagen como el proceso mediante el cual se definen las clases o categorías presentes en la imagen y cada píxel que parte de la misma es asignado a una de estas categorías. Sheikholeslami et al [86] plantea que el objetivo de la clasificación de imágenes geográficas consiste en la categorización de imágenes de acuerdo al tipo de cobertura o uso del suelo, por ejemplo para distinguir entre áreas forestales y residenciales.

Chuvieco [87] trata de forma detallada la clasificación digital, partiendo de la consideración de que es posible abordar este proceso teniendo en cuenta únicamente los niveles digitales de la imagen. En este contexto se entiende por nivel digital, un valor numérico que define el nivel de gris aplicado a cada píxel durante la visualización de la imagen, dicho valor corresponde a la medición de la radiancia efectuada por el sensor en el momento de adquirir la imagen. El proceso de clasificación se componen de tres fases: entrenamiento, asignación, comprobación y verificación de resultados según Chuvieco [87]:

- ✓ **Entrenamiento:** Consiste en la definición de las categorías.
- ✓ **Asignación:** Consiste en el agrupamiento de los píxeles en las categorías definidas. El problema en la fase de asignación está en definir qué criterio aplicar para el agrupamiento de los píxeles, es decir, se debe establecer cuando se considera que el nivel digital de un píxel es similar u homogéneo a otro. Algunos métodos de clasificación usan solo la información espectral (píxel por píxel), y otros buscan mejorar los resultados utilizando las relaciones existentes entre cada píxel y sus vecinos según Sheikholeslami et al [86].
- ✓ **Comprobación y verificación de resultados:** Tradicionalmente se ha aceptado la división de los métodos de clasificación en dos grupos según Chuvieco [87]:
 - **Supervisado:** El interprete define zonas (grupos de píxeles) representativas de una clase o categoría, basándose en un conocimiento previo de la zona de estudio. Más adelante estas definiciones son utilizadas para clasificar cada píxel en función de la similitud de sus niveles digitales con los de los píxeles utilizados como referencia. Las clases resultantes son llamadas también clases informacionales.
 - **No supervisado:** No se requiere un conocimiento previo de la zona de estudio. Este método desarrolla una clasificación automática mediante el agrupamiento de valores homogéneos dentro de la imagen. Las clases resultantes son llamadas también espectrales.

En la tabla 1 se relacionan algunas técnicas de reconocimiento de patrones, aplicados a la clasificación de imágenes, las cuales se encuentran brevemente descritas en Tso et al [88].

Tabla 1. Técnicas de reconocimiento de patrones, aplicadas a la clasificación de imágenes según [88].

No Supervisadas	Algoritmos de agrupamiento con migración de las medias (o centros)	
	ISODATA	
	Refinamiento de agrupación	
	Agrupamiento difuso	
	Detección Vacuum shell	
Supervisadas	Métodos estadísticos	Paralelepípedos
		Mínima Distancia
		Máxima probabilidad
	Redes Neuronales artificiales	
	Basadas en el conocimiento	Árboles de decisión
		Extracción automática de reglas difusas

Además de las anteriores técnicas, se han explorado otros métodos que combinan los resultados obtenidos usando diferentes reglas de decisión, como reglas de elección, redes bayesianas según Grossman [89], razonamiento probatorio, múltiples enfoques de redes neuronales según Cetin et al [90], autómatas celulares según Rosin [91] y algoritmos evolutivos como los propuestos por Quirin et al [92], Larch [93], Lee et al [94] y Lozano et al [95].

De igual forma en la literatura se encuentran otras propuestas alternativas a las técnicas y algoritmos ya mencionados. Es el caso propuesto por Vatsavai et al [96] donde se presenta un conjunto de nuevos algoritmos de clasificación, agrupados bajo el nombre de *Miner, que fueron investigados y desarrollados como parte del proyecto TerraSIP financiado por la NASA. Estos nuevos algoritmos son resultado de la combinación de otras técnicas. Por ejemplo, el Clasificador espectral, espacial y temporal basado en el conocimiento SSTKC (Spectral, Spatial, and Temporal Knowledge-based Classifier), una fusión de sistemas basados en el conocimiento y el clasificador de máxima probabilidad, dirigido a la clasificación de imágenes multi-espectrales utilizando conocimiento de bases de datos espaciales auxiliares. El sistema de clasificación híbrido HCS (Hybrid Classification System), tiene por objetivo combinar métodos estadísticos de reconocimiento de patrones con sistemas de clasificación basados en el conocimiento. Los clasificadores espaciales o contextuales, buscan incorporar la dependencia espacial a los problemas de clasificación. Se distinguen en este grupo dos modelos, los campos aleatorios de Markov MRF (Markov Random Field) y la autoregresión espacial SAR (Spatial Auto-regression).

4.3.1 Utilización de algoritmos genéticos para la clasificación de imágenes

Gabriel et al [85] explica que los algoritmos genéticos son algoritmos de búsquedas inspirados en los mecanismos de selección natural de las especies y la combinación genética que se presenta en la reproducción de individuos. Estos algoritmos utilizan una estructura de datos simple llamada cromosoma para representar posibles soluciones a un problema específico, y aplica a esas estructuras diferentes operadores y combinaciones de ellos de forma que la información importante sea preservada.

En relación con el empleo de estos algoritmos en el proceso de clasificación de imágenes, se encuentran tres enfoques. El primero se refiere a la clasificación de estos algoritmos directamente al proceso de clasificación de imágenes no clasificadas. El segundo enfoque plantea la aplicación de esta técnica en la extracción (aprendizaje) de reglas de clasificación sobre imágenes previamente clasificadas por expertos. Un tercer enfoque utiliza los algoritmos genéticos para generar un nuevo algoritmo clasificador a partir de la combinación de operadores de bajo nivel aplicables a las imágenes.

También encontramos varios trabajos sobre clustering con estrategias evolutivas como son los propuestos por Lee et al [94] y Lozano et al [95]. El clustering es uno de los mecanismos más utilizados en el proceso de clasificación, y consiste en el agrupamiento de píxeles con valores digitales similares.

5. Métodos automáticos de anotación semántica de imágenes de teledetección espacial

Partiendo de la literatura estudiada se puede inferir que los métodos de anotación semántica de imágenes satelitales estudiados proceden de igual manera que los métodos tradicionales de anotación semántica de imágenes, incluso estos métodos se acoplan fácilmente al modelo de anotación propuesto por Zhang [12], partiendo de un conjunto de imágenes de entrenamiento anotadas previamente, introducidas por cada concepto que proporciona el usuario, se procede a la extracción de los rasgos más significativos de cada una de estas imágenes, luego se utiliza un modelo probabilístico de generación de rasgos y con este paso tenemos implementado un modelo de anotación de imagen según Zhang [12], cuando desean anotar una nueva imagen, proceden a la extracción de los rasgos de la misma y aplican el modelo de anotación ya implementado y como resultado obtienen la misma imagen de entrada pero ya anotada.

Los métodos de anotación semántica de imágenes de teledetección espacial encontrados en la literatura utilizan para realizar automáticamente el proceso de anotación el modelo Latent Dirichlet Allocation conjuntamente con un método de clasificación supervisada como Maximum Likelihood y también encontramos otro modo de proceder que automatiza el proceso de anotación semántica mediante el uso del método SIFT (Scale Invariante Feature Transform) en combinación con el algoritmo Expectation-Maximization y el criterio de Minimum Description Length.

Estos métodos analizados realizan el proceso de anotación semántica partiendo de conceptos introducidos por el usuario, y por cada concepto introducido por el usuario se proporcionan además un conjunto de imágenes de entrenamiento, en ninguno de los casos encontrados hacen uso de las ontologías como forma de representación formal del conocimiento para la base semántica de la anotación. En el caso del uso del modelo LDA sucedieron varios errores en el proceso de clasificación debido a que las imágenes que se utilizaron en el conjunto de prueba contenían más de cinco clases que no pudieron ser encontradas debido a que las mismas no se proporcionaron en las imágenes del conjunto de entrenamiento, estos errores en el proceso de clasificación surgen debido al único empleo de la media y la desviación estándar, siendo necesario definir conceptos más amplios que cubran las clases que no fueron identificadas durante el proceso de clasificación. También es necesario resaltar el uso de otros rasgos característicos en las imágenes como la textura y/o el trabajo con imágenes multi-espectrales con el objetivo de mejorar el funcionamiento del modelo LDA y de este modo mejorar el proceso de anotación. En el segundo caso, haciendo uso del método SIFT es necesario resaltar también que sucedieron errores durante el proceso de anotación debido a que también se encontraron clases contenidas en las imágenes de prueba que no fueron definidas en las imágenes pertenecientes al conjunto de entrenamiento.

6. Anotaciones semánticas de información geográfica

Oliva [97] hace una propuesta de metodología para la anotación semántica de datos geográficos compuesta por tres componentes: modelo de anotación semántica, arquitectura de persistencia de las anotaciones semánticas y procedimiento de anotación.

Según Oliva [97] el modelo de anotación semántica define qué se identificará como anotación semántica y sentará las bases para el resto de los componentes de la metodología. Se define anotación semántica de datos geográficos como el establecimiento de vínculos, asociaciones o enlaces entre los objetos geográficos representados en el dato y el conocimiento previamente formalizado y estos vínculos deben ser susceptibles de ser procesados en un ordenador y ser recuperados. Se explica que las geo-ontologías son estructuras en las que es posible la integración o asociación dato-conocimiento. Se explica también que las anotaciones de los objetos geográficos se modelarán como instancias de conceptos pertenecientes a geo-ontologías, de modo que hay que distinguir de cual ontología se está haciendo referencia. Oliva [97] explica que un modelo de anotación semántica de datos geográficos debe tener acceso a mecanismos para la persistencia de las anotaciones semánticas, es decir debe valerse de alguna arquitectura que garantice la recuperación, almacenamiento y gestión de las anotaciones semánticas, así como posibilitar la recuperación y descubrimiento de la información a partir del conocimiento y del conocimiento partiendo de la información espacial.

La arquitectura de persistencia de las anotaciones semánticas abarca la consideración de todo lo relacionado con el almacenamiento y recuperación de las anotaciones semánticas: estructuración e interpretación de la información y el conocimiento almacenado, costo computacional de la gestión de las anotaciones así como la capacidad de representar la realidad anotada según Oliva [97] y en este mismo trabajo se propone una arquitectura basada en ontologías para la persistencia de las anotaciones semánticas de datos geográficos como se muestra en la figura 12.

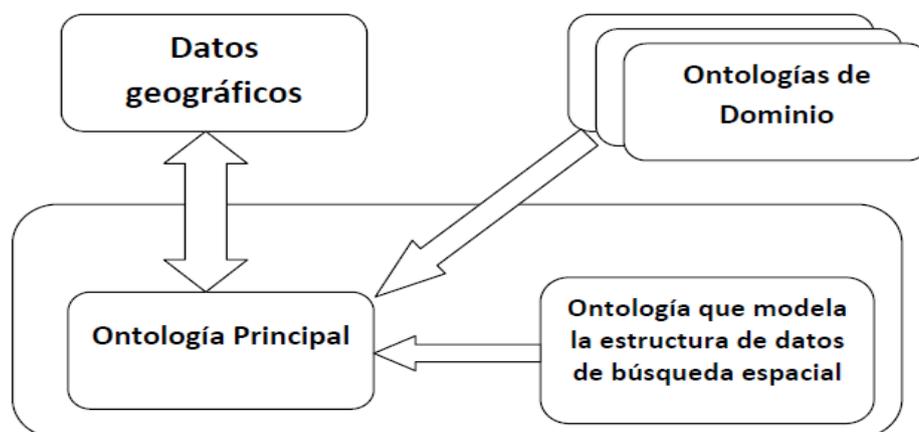


Fig.12. Componentes de la arquitectura de persistencia de anotaciones semánticas según [97]

Esta arquitectura tiene como componente fundamental una ontología principal que servirá de soporte a las anotaciones, esta ontología estará organizada en forma de alguna estructura de datos de búsqueda espacial de modo que la recuperación de la información y conocimiento almacenado sea efectiva. La estructura de datos de búsqueda espacial, en la que estarán organizadas las anotaciones, debe estar conceptualizada en una ontología y cada anotación que se quiera establecer constituirá una instancia de algún concepto previamente formalizado, la generalidad deben ser conceptos provenientes de ontologías de dominio. La arquitectura que propone se basa en la correspondencia entre los objetos geográficos representados en el dato e instancias que se formalizarán en la ontología principal. Esta

correspondencia es posible sólo si los conceptos de los cuales son instancias los objetos geográficos a modelar existen previamente, en caso de no ser así será necesario crearlos en la propia ontología o en alguna otra que se esté reutilizando. Debe garantizar que los especialistas que realicen la anotación puedan valerse de conocimiento previamente formalizado, así como que tengan la posibilidad de formalizar nuevos conocimientos. Esta arquitectura debe posibilitar la reutilización del conocimiento representado formalmente con anterioridad en ontologías de dominio, por lo que la ontología principal incluye información sobre cuáles ontologías se van a citar y la forma en la que se puede acceder a ella. Es importante que la arquitectura de persistencia permita recuperar los conceptos relacionados con un objeto geográfico a un costo de procesamiento computacional aceptable según Oliva [97].

Los procedimientos de anotación son los mecanismos mediante los cuales se realiza la vinculación entre los objetos geográficos presentes en el dato y los conocimientos previamente formalizados. Los procedimientos de anotación semántica de datos geográficos dependen en gran medida del tipo o formato de dato que se desee anotar así como del factor de enlace dato-conocimiento que se seleccione según Oliva [97].

En Klien [98] se propone un modelo conceptual para el procedimiento de anotación semántica de datos geográficos en entornos de servicios web, además de desarrollar un método para apoyar automáticamente el proceso de anotación que evalúa la validez de anotaciones existentes y sugiere posibles nuevas anotaciones. Explican que el uso de las ontologías en los sistemas de información para valorar la interoperabilidad semántica entre fuentes de datos geoespaciales está lejos de ser un proceso consistente; y que mientras más se desarrollan ontologías de dominio geoespacial se observa que no existe una estrategia bien definida para el proceso de asociación de las descripciones de categorías existentes en las ontologías con los rasgos geográficos de una fuente de datos en particular. De esta manera identifican la necesidad de definir un modelo conceptual y una estrategia para el proceso de anotación que posibilite la creación de anotaciones consistentes y reproducibles. Explican que el problema de generar anotaciones semánticas ha sido reconocido como uno de los más serios obstáculos para la realización de la Web Semántica Geoespacial y partiendo de esto presentan una estrategia basada en reglas para la anotación semántica de geodatos que combina la Web Semántica y las tecnologías de los Servicios Web Geoespaciales. Estas reglas definen condiciones para la identificación de conceptos geoespaciales. Se hace referencia al rol de las relaciones espaciales en el proceso de automatización de la anotación semántica de geodatos. Las relaciones espaciales entre las entidades son características de la información geográfica y a menudo son tan importantes como las propias entidades. En las ontologías de dominio geoespacial, las relaciones taxonómicas y no taxonómicas son utilizadas para definir conceptos del mundo físico y para diferenciarlos entre ellos. Las relaciones espaciales juegan un importante rol para la definición e identificación de conceptos espaciales. En el dominio geoespacial, las relaciones entre entidades espaciales son tan importantes como las propias entidades. Las relaciones espaciales han sido clasificadas de varias maneras y se hace referencia a la clasificación de estas relaciones espaciales basándose en su comportamiento característico en el espacio; partiendo de esto presentan la siguiente clasificación: *relaciones topológicas* que hacen referencia a propiedades tales como conectividad, adyacencia e intersección entre entidades geoespaciales, este tipo de relación se mantiene invariante bajo transformaciones topológicas consistentes, tales como rotación, traslación y escala, *relaciones direccionales* que tratan con el orden en el espacio y están basadas en la existencia de un vector de espacio y, por consiguiente están sujetas a cambios bajo la rotación, mientras se mantienen invariantes a la traslación y escala del marco de referencia y por último tenemos las *relaciones de distancia* que hacen referencia a las distancias geográficas entre los objetos geoespaciales, estas reflejan el concepto de métrica, de esta manera cambian bajo los efectos de escala pero se comportan invariantes frente a la traslación y rotación.

Según Klien [98] las descripciones de categorías formales en las ontologías pueden ser usadas para explicar información. Podemos nombrar a esta explicación proceso de anotación

semántica. Con la anotación semántica se hace posible asociar múltiples conceptualizaciones al contenido de una fuente de datos y de esta manera es más fácil encontrarla en diferentes contextos. La disponibilidad de las anotaciones semánticas para la información geográfica puede extender el alcance para valorar la interoperabilidad semántica considerablemente en los entornos de los servicios web de información geográfica. Los enfoques para la valoración de la interoperabilidad semántica que van más allá del análisis de texto y hacer uso de las fórmulas lógicas desde ontologías ha sido aplicado exitosamente. Generar anotaciones con ontologías es un proceso difícil y que consume tiempo y actualmente no existen estrategias comunes para indicar como anotar fuentes de información geográfica en entornos de servicios web.

7. Vinculando teledetección, anotaciones semánticas, ontologías a sistemas de información geográfica

Los Sistemas de Información Geográfica son programas que proporcionan, de forma modular o integrada, herramientas para la introducción, almacenamiento, gestión, análisis y representación gráfica de datos que se caracterizan por la existencia de una componente espacial, por tanto se trata de datos georreferenciados.

7.1 Teledetección y sistemas de información geográfica

La combinación de la teledetección y los Sistemas de Información Geográfica permiten hacer frente y apoyar la toma de decisiones como una metodología efectiva y poderosa en diferentes problemas presentes en nuestra sociedad. Estas herramientas informativas y de análisis visual han sido ampliamente aplicadas a mejorar el monitoreo y conservación de diferentes recursos y entornos naturales. Diversos trabajos demuestran que la vinculación de ambas posibilita la solución o el estudio de problemas como los encontrados en la literatura estudiada como los presentados por Angueira et al [99], Degioanni et al [100], Castaño [101], García et al [102].

En Basso et al [103] se explica que la tecnología de la percepción remota no estaría completa sin el surgimiento de los Sistemas de Información Geográfica; que esta percepción remota representa una tecnología para la adquisición sinóptica de datos espaciales y la extracción de información específica de una escena, y que durante los últimos años ha crecido exponencialmente la demanda de la percepción remota como fuente de entrada de datos para el desarrollo de bases de datos espaciales; estas informaciones obtenidas mediante la teledetección son particularmente interesantes para el desarrollo de bases de datos de Sistemas de Información Geográfica.

7.2 Ontologías en los sistemas de información geográfica

La posibilidad de integración de la Información Geográfica (IG) se hace cada vez mayor, debido al incremento de la conectividad mundial y del aumento de la disponibilidad de información geográfica a nivel mundial. Las presentes y futuras generaciones de Sistemas de Información Geográfica (SIG) están sustentadas en la creación de la colección y especificación sistemática de entidades geográficas, sus propiedades y sus relaciones. En este caso las ontologías juegan un importante rol en la construcción de los SIG, debido a que permiten el establecimiento de correspondencias e interrelaciones entre diferentes dominios de las entidades espaciales según Renato [104].

En Fonseca et al [105] se explica que debemos considerar que las geo-ontologías deberían ser componentes esenciales del universo lógico para el modelamiento de datos geográficos. Se presenta un énfasis especial a la utilización de estructuras ontológicas para la

integración de la información semántica entre Sistemas de Información Geográfica y Sistemas de Percepción Remota. Enfocándose en el hecho de que, con el desarrollo de nuevos sensores y nuevas estrategias de colección de datos se está incrementando la necesidad de la implementación de nuevas arquitecturas que posibiliten la integración de la información geográfica, presentan la implementación de una arquitectura para un Sistema de Información Geográfica Gobernado por Ontologías (Ontology-Driven Geographic Information System).

Las ontologías representan un rol esencial en la construcción de un Sistema de Información Geográfica, dado que posibilitan el establecimiento de correspondencias e interrelaciones entre los diferentes dominios de las entidades espaciales y sus relaciones según Fonseca et al [105].

Los Sistemas de Información Geográfica Gobernados por Ontologías (SIGGO) son implementados haciendo uso de componentes de software derivados desde varias ontologías. Estos componentes de software son clases que pueden ser utilizadas para el desarrollo de nuevas aplicaciones. La estructura de un SIGGO tiene dos principales aspectos: *la generación de conocimiento y el uso del conocimiento*. La *generación de conocimiento* implica la especificación de ontologías haciendo uso de un editor de ontologías, la generación de nuevas ontologías partiendo de ontologías existentes y la traducción de estas ontologías en componentes de software. El *uso del conocimiento* se basa en la utilización de estas ontologías disponibles para los usuarios y los metadatos de información acerca de la información disponible que proporcionan dichas ontologías. Un conjunto de clases que contiene datos y operaciones constituye la funcionalidad del sistema. Estas clases están vinculadas a fuentes de información geográfica a través del uso de mediadores según Fonseca et al [105].

Los componentes principales de la arquitectura de un SIGGO son los siguientes según Fonseca et al [105]:

- ✓ **El servidor de ontologías:** este servidor de ontologías representa el rol fundamental en un SIGGO debido a que proporciona la conexión entre todos los componentes principales. Este servidor es también responsable de hacer disponibles las ontologías para las aplicaciones, la conexión con la fuente de información se realiza a través de mediadores. Los mediadores representan piezas de software con conocimiento embebido.
- ✓ **Las ontologías:** están representadas por dos tipos de estructuras, las especificaciones y las clases. Las especificaciones están hechas por expertos y almacenadas de acuerdo a sus rasgos distintivos (partes, funciones y atributos) y sus interrelaciones semánticas. Esta estructura proporciona información acerca del significado de la información disponible. Las clases representan el resultado de la traducción de las ontologías. Son componentes de software que pueden ser usados para desarrollar nuevas aplicaciones.
- ✓ **Las fuentes de información:** las fuentes de información geográfica en un SIGGO pueden ser cualquier tipo de base de datos geográfica siempre y cuando se comprometan con un mediador, el cual tiene la funcionalidad de extraer la información necesaria para generar una instancia de una entidad de una ontología y también puede incorporar nueva información en caso de una actualización.
- ✓ **Las aplicaciones:** una de las aplicaciones de un SIGGO es la recuperación de la información. Los mediadores proporcionan instancias de las entidades disponibles en el servidor de ontologías. Los usuarios pueden explorar la información en diferentes niveles de detalles dependiendo el nivel de ontología usado. Otro tipo de aplicaciones puede ser el desarrollo, tal como actualización de una base de datos y diferentes tipos de procesamiento de datos geográficos, incluyendo análisis estadístico y procesamiento de imagen.

En Fonseca et al [105] se explica que en la arquitectura de un SIGGO existen diferentes niveles de ontologías y como consecuencias existen también diferentes niveles de detalles de

información. Las ontologías de bajo nivel corresponden a información bien detallada mientras que las ontologías de alto nivel corresponden a información más general.

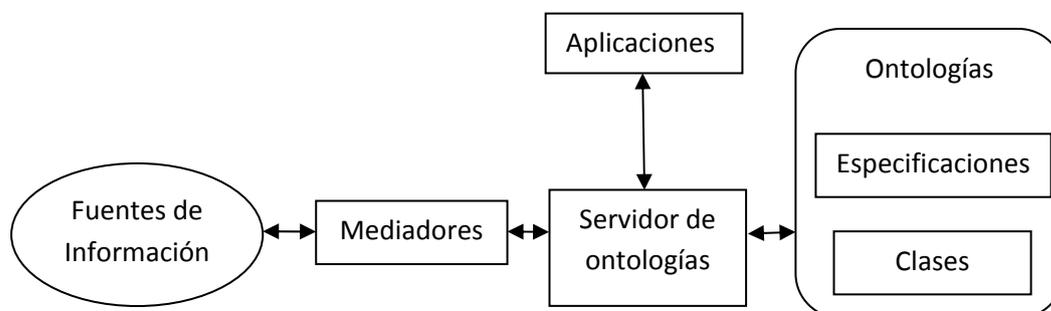


Fig.13. Arquitectura básica de un SIGGO según [105]

7.3 Anotaciones semánticas en sistemas de información geográfica

Según Oliva et al [106] el desarrollo futuro de los Sistemas de Información Geográfica está orientado a la inclusión de semántica en la gestión y mencionan que un primer paso para la inclusión de semántica en la gestión de estos SIG es la asociación o anotación semántica de los datos espaciales, es decir la correspondencia entre los objetos geográficos presentes en el dato espacial y conocimiento previamente formalizado en geo-ontologías.

La interoperabilidad semántica de la información y los servicios geográficos se ha convertido en una de las principales líneas de investigación de las Ciencias de la Información Geográfica según Raubal [107]. La interoperabilidad semántica puede lograrse mediante el uso de geo-ontologías de dominio y anotaciones semánticas de datos espaciales que utilicen dichas geo-ontologías según Klien [108].

Es necesaria la asociación dato-conocimiento espacial para lograr una mayor interoperabilidad y dotar a los actuales SIG de herramientas de gestión de estas anotaciones semánticas Oliva et al [106].

8. Conclusiones

El término auto-anotación semántica de imágenes es un término bastante flexible que posibilita cómodamente la realización del proceso natural de recuperación de imágenes desde una base de datos o desde un repositorio de datos para las comunidades de usuarios. En este documento se realizó un estudio acerca de los métodos automáticos de anotación semántica de imágenes, así como de las técnicas de procesamiento de imágenes que se deben realizar para llevar a cabo este proceso de anotación. Los métodos encontrados en la literatura estudiada proceden a anotar imágenes de teledetección de manera semántica partiendo de conceptos introducidos por el usuario; en ninguno de los métodos analizados se proceden a la utilización de ontologías de dominio como representación formal de conocimiento, lo que ha demostrado ser la vía más eficiente y eficaz para realizar recuperaciones de grandes volúmenes de imágenes.

Una posible vía de solución a esta limitante puede estar en el desarrollo de un método automático de anotación semántica de imágenes de teledetección basados en ontologías mediante la combinación de métodos de extracción de rasgos de las imágenes y la

generación de modelos como el LDA y su ampliación y una arquitectura de persistencia de anotaciones semánticas de datos geográficos y su integración a Sistemas de Información Geográfica Gobernados por Ontologías.

Referencias bibliográficas

1. Jean-Baptiste Bordes, Henri Maître, (2007). "Semantic annotation of satellite images".
2. <http://siga.cna.gob.mx/SIGA/Diccionarios/glosario.htm>. accedido 22-09-2009, 2009.
3. <http://es.wikipedia.org/wiki/Sem%C3%A1ntica>. accedido 22-09-2009, 2009.
4. Gruber, T. R., (1993). "Toward Principles for the Design of Ontologies Used for Knowledge Sharing." International Journal of Human and Computer Studies, 43 (5-6), 1995. Pp. 907-928 (también disponible como Technical Report KSL 93-04, Knowledge Systems Laboratory, Stanford University).
5. Borst, W., (1997), "Construction of Engineering Ontologies for Knowledge Sharing and Reuse". Ph.D. Dissertation, University of Twente.
6. GUARINO, N, (1998). "Formal Ontology and Information Systems". Proceedings of FOIS '98. National Research Council, LADSEB-CNR. Disponible en <http://citeseer.ist.psu.edu/guarino98formal.html> (Fecha de acceso: 29 de Abril de 2005).
7. Barry Smith and David M. Mark, (2001). "Geographic Categories: An Ontological Investigation", International Journal of Geographical Information Science, 15 (7), 591-612.
8. Klien Eva and Fitzner Daniel I. and Maué Patrick, (2007). "Baseline for Registering and Annotating Geodata in a Semantic Web Service Framework".
9. Grcar, M. and E. Klien, (2007). "Using Term-matching Algorithms for the Annotation of Geoservices.", In Proceedings of Web Mining 2.0 Workshop. In conjunction with ECML-PKDD 2007, Warsaw, Poland. <http://www.swingproject.org/publications/document/100/>, (last accessed 15/08/2008).
10. http://en.wikipedia.org/wiki/Automatic_image_annotation. 24-09-2009, 2009.
11. Carlos Pérez Gutiérrez, Angel Luis Muñoz Nieto, (2002). "Teledetección Nociones y Aplicaciones".
12. Yu-Jin Zhang, (2007). "Semantic-Based Visual Information Retrieval.", Tsinghua University, Beijing, China.
13. Chang, E., Goh, K., Sychay, G., & Wu, G, (2003). "CBSA: Content-based soft annotation for multimodal image retrieval using Bayes Point Machines.", IEEE Transactions on Circuits and Systems for Video Technology: Special Issue on Conceptual and Dynamic Aspects of Multimedia Content Description, 13(1), 28-38.
14. Herbrich, R., Graepel, T., & Campbell, C, (2001). "Bayes Point Machines.", Journal of Machines Learning Research, 1, 245-279.
15. http://images.google.com/cu/imgres?imgurl=http://research.microsoft.com/en-us/um/cambridge/projects/infernet/docs/Multiclass_BPM.jpg&imgrefurl=http://research.microsoft.com/en-us/um/cambridge/projects/infernet/docs/Multiclass%2520classification.aspx&usq=h6RRFACfY1PKvtKUh8HWRn7828=&h=399&w=718&sz=79&hl=es&start=1&tbnid=Am_pOipONrr2xM:&tbnh=78&tbnw=140&prev=/images%3Fq%3DBayesian%2BPoint%2BMachines%26gbv%3D2%26hl%3Des%26sa%3DG.
16. Feng, H., Shi, R., & Chua, T, (2004). "A bootstrapping framework for annotation and retrieving WWW images.", In Proceedings of the ACM Conference on Multimedia, New York (pp. 960-967).
17. Vapnik, V, (1995). "The nature of statistical learning theory". New York: Springer-Verlag.
18. http://en.wikipedia.org/wiki/Support_vector_machine. 28-09-2009, 2009.
19. http://catarina.udlap.mx/u_dl_a/tales/documentos/mcc/sanchez_1_se/capitulo1.pdf. 28-09-2009, 2009.
20. Diego Lizarazo Arias, (2007). "Semántica de las imágenes: figuración, fantasía e iconicidad".
21. Jianye Ye, Xiangdong Zhou, Jian Pei, Lian Chen, Liang Zhang, (2005). "A Stratification-Based Approach to Accurate and Fast Image Annotation.", WAIM 2005, LNCS 3739, pp. 284-296.
22. Heidy Marin-Castro, Enrique Sucar, and Eduardo Morales,(2007). "Automatic Image Annotation Using a Semi-supervised Ensemble of Classifier.", CIARP 2007, LNCS 4756, pp. 487-495.
23. Yong Wang, TaoMei, Shaogang Gong, Xian-Sheng Hua, (2008). "Combining global, regional and contextual features for automatic image annotation."
24. Tatiana Tommasi, Francesco Orabona, Barbara Caputo, (2008). "Discriminative cue integration for medical image annotation."
25. Lokesh Setia, Alexandra Teynor, Alaa Halawani, Hans Burkhardt, (2008). "Grayscale medical image annotation using local relational features."
26. Wei Li and Maosong Sun, (2006). "Incorporating Prior Knowledge into Multi-label Boosting for Cross-Modal Image Annotation and Retrieval.", AIRS 2006, LNCS 4182, pp. 404 - 415.
27. Carlos Hernández-Gracidas and L. Enrique Sucar, (2007). "Markov Random Fields and Spatial Information to Improve Automatic Image Annotation.", PSIVT 2007, LNCS 4872, pp. 879-892.
28. Alexei Yavlinsky, Edward Schofield, and Stefan Ruger, (2005). "Automated Image Annotation Using Global Features and Robust Nonparametric Density Estimation.", CIVR 2005, LNCS 3568, pp. 507-517.

29. Xiaojun Qi, Yutao Han, (2007). "Incorporating multiple SVMs for automatic image annotation."
30. Amalia Kallergi, Yun Bei, Fons J. Verbeek, (2009). "The Ontology Viewer: Facilitating Image Annotation with Ontology Terms in the CSIDx Imaging Database.", section Imaging & BioInformatics - Imagery & Media group Leiden Institute of Advanced Computer Science (LIACS), Leiden University. Niels Bohrweg 1, 2333 CA Leiden, The Netherlands, Workshop on Visual Interfaces to the Social and the Semantic Web (VISSW2009), IUI2009, Sanibel Island, Florida, USA.
31. Chin-Hui Lee, Rui Shi, Tat Seng Chua, (2007). "Enhancing image annotation by integrating concept ontology and text-based bayesian learning model.", International Multimedia Conference Proceedings of the 15th international conference on Multimedia.
32. Hollink, L., Schreiber, A., Wielemaker, J., & Wielinga, B (2003). "Semantic annotation of image collections.", In Proceedings of the Workshop on Knowledge Markup and Semantic Annotation, Sanibel Island, FL.
33. Dominic Mainz, Katrin Weller, Jürgen Mainz, (2008). "Semantic Image Annotation and Retrieval with IKen.", Proceedings of the Poster and Demonstration Session of the Seventh International Semantic Web Conference (ISWC2008), CEUR volume 401, Karlsruhe, Germany, October 28.
34. Yohan Jin, Lei Wang and Latifur Khan, (2005). "Improving Image Annotations Using WordNet.", MIS 2005, LNCS 3665, pp. 115–130.
35. Manuel Moller, Sven Regel, and Michael Sintek, (2009). "RadSem: Semantic Annotation and Retrieval for Medical Images.", ESWC 2009, LNCS 5554, pp. 21–35.
36. Lei Shi, Guochang Gu, Haibo Liu, Jing Shen, Lei Shi, (2008). "A Semantic Annotation Algorithm Based on Image Regional Object Ontology.", csse, vol. 4, pp.540-543, International Conference on Computer Science and Software Engineering.
37. Kosmas Petridis, Frederic Precioso, Thanos Athanasiadis, Yannis Avrithis and Yiannis Kompatsiaris, (2005). "Combined Domain Specific and Multimedia Ontologies for Image Understanding."
38. Manuel Möller, Saikat Mukherjee, (2009). "Context-driven ontological annotations in DICOM images: Towards Semantic Pacs."
39. Zhiguo Gong, Qian Liu, and Jingbai Zhang, (2006). "Automatic Image Annotation by Mining the Web.", DaWaK 2006, LNCS 4081, pp. 449–458.
40. Hongtao Xu, Xiangdong Zhou, Lan Lin, Yu Xiang, and Baile Shi, (2009). "Automatic Web Image Annotation via Web-Scale Image Semantic Space Learning.", APWeb/WAIM 2009, LNCS 5446, pp. 211–222.
41. Roger C.F. Wong, Clement H.C. Leung, (2008). "Automatic Semantic Annotation of Real-World Web Images.", IEEE Transaction on Pattern Analysis and Machine Intelligence, pp. 1933-1944, (vol. 30 no. 11).
42. Roberto Basili, Riccardo Petitti and Dario Saracino, (2007). "Mining Web Data for Image Semantic Annotation.", AI*IA 2007, LNAI 4733, pp. 674–685.
43. Andreas Walter and Gabor Nagypal, (2008). "The Combination of Techniques for Automatic Semantic Image Annotation Generation in the IMAGINATION Application.", ESWC 2008, LNCS 5021, pp. 879–883.
44. Town, C. (2004). "Ontology based visual information processing.", [doctoral dissertation].Cambridge, UK: University of Cambridge.
45. Hu, B., Dasmahapatra, S., Lewis, P., & Shadbolt, N. (2003). "Ontology-based medical image annotation with description logics.", In Proceedings of the IEEE Conference on Tools with Artificial Intelligence, Sacramento, CA (p. 77).
46. Soo, V., Lee, C., Li, C., Chen, S., & Chen, C. (2003). "Automatic semantic annotation and retrieval based on sharable ontology and case-based learning techniques.", In Proceedings of the Joint Conference on Digital Libraries, Houston, TX (61–72).
47. Shen, H., Ooi, B., & Tan, K. (2000). "Giving meanings to WWW images.", In Proceedings of the ACM Conference on Multimedia, Marina del Rey, CA (pp. 39–47).
48. Lieberman, H., Rosenzweig, E., & Singh, P. (2001). "Aria: An agent for annotating and retrieving images.", IEEE Computer, 34(7), 57–61.
49. Carson, C., Thomas, M., Belongie, S., Hellerstein, J., & Malik, J. (1999). "Blobworld: A system for region-based image indexing and retrieval.", Proceedings of the Conference on Visual Information Systems (pp. 509–516). Amsterdam, Netherlands.
50. Shi, J., & Malik, J. (2000). "Normalized cuts and image segmentation.", IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8), 888–905.
51. Dietterich, T., Lathrop, R., & Lozano-Perez, T. (1997). "Solving the multiple instance learning with axis-parallel rectangles.", Artificial Intelligence, 89(1-2), 31–71.
52. Jun Yang, "Review of Multi-Instance Learning and Its applications.", <http://www.cs.cmu.edu/~juny/MILL/review.htm>. 29-10-2009, 2009.
53. Moon, T. (1996). "The expectation-maximization algorithm.", IEEE Signal Processing, 13(6), 47–60.
54. Mori, Y., Takahashi, H., & Oka, R. (1999). "Image-to-word transformation based on dividing and vector quantizing images with words.", In Proceedings of the Workshop on Multimedia Intelligent Storage and Retrieval Management, Orlando, FL.

55. Duygulu, P., Barnard, K., Freitas, N., & Forsyth, D. (2002). "Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary.", In Proceedings of the European Conference on Computer Vision, Copenhagen, Denmark (pp. 97–112).
56. Barnard, K., & Forsyth, D. (2001). "Learning the semantics of words and pictures.", In Proceedings of the Conference on computer vision, Vancouver, Canada (pp. 408–415).
57. Jeon, J., Lavrenko, V., & Manmatha, R. (2003). "Automatic image annotation and retrieval using cross-media relevance models.", In Proceedings of the ACM Conference on SIGIR, Toronto, Canada (pp. 119–126).
58. Li, J., Gray, R., & Olshen, R. A. (2000). "Multiresolution image classification by hierarchical modeling with two dimensional hidden markov models.", *IEEE Transactions on Information Theory*, 34(5), 1826–1841.
59. Hofmann, T. (1998). "Learning and representing topic: A hierarchical mixture model for word occurrence in document databases.", In Proceedings of the Workshop on Learning from Text and Web, Pittsburgh, PA.
60. Lavrenko, V. y Croft, W. B. (2001). "Relevance-Based Language Models.", Proceedings of the ACM SIGIR 01 conference, pp. 120-127.
61. Li, J., & Wang, J. (2003). "Automatic linguistic indexing of pictures by a statistical modeling approach.", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9), 1075–1088.
62. Jacek Koronacki, Zbigniew W. Ra's, S awomir T. Wierzcho'n y Janusz Kacprzyk, (2010). "Advances in Machine Learning II".
63. Tsai, C.-F., Hung, C. (2008). "Automatically annotating images with keywords: A review of image annotation systems", *Recent Patents on Computer Science*.
64. Tsai C-F, McGarry K, Tait JC. (2006) "A modular support vector image indexing and classification system". *ACM Transactions on Information Systems a*; 24(3): 353-379.
65. Mitchell T. (1997). "Machine Learning". McGraw Hill, New York.
66. Bishop CM. (1995). "Neural networks for pattern recognition". Oxford University Press, Oxford.
67. Ghosh, J. (2002). "Multiclassifier systems: back to the future". Proceedings of the 3rd International Workshop on Multiple Classifier Systems, Cagliari, Italy, June 24-26.; 1-15.
68. Frosyniotis D, Stafylopatis A, Likas A, (2003). "A divide-and-conquer method for multi-net classifiers". *J Patt Anal Appl*; 6(1): 32-40.
69. Ho TK, Hull JJ, Srihari SN. (1994). "Decision combination in multiple classifier systems". *IEEE Trans Patt Anal Mach Intel*; 16(1): 66-75.
70. Duin RPW, Kittler J, Hater M, Mates J. (1998). "On combining classifiers". *IEEE Trans Patt Anal Mach Intel*; 20(3): 226-239.
71. Haykin S. (1999). "Neural networks: a comprehensive foundation", 2nd Edition. Prentice Hall, New Jersey.
72. Breiman, L. (1996). "Bagging predictors". *Machine Learning* 1996; 24(2): 123-140.
73. Schapire RE. (1990). "The strength of weak learnability". *Mach Learn*; 5(2): 197-227.
74. Wolpert DH. (1992). "Stacked generalization". *Neural Networks*; 5(2): 241-259.
75. Jain AK, Murty MN, Flynn PJ, (1999). "Data clustering: a review". *ACM Comp Surv*; 31(3): 264-323.
76. Eakins JP, (1997). "Automatic image content retrieval – how usable is the technology". *Int J Elect Lib Res*; 7(1): 63-88.
77. Eakins JP, Graham M, (1999). "Content-based image retrieval". A Report to the JISC Technology Applications Programme, University of Northumbria at Newcastle. [On-line]. Available from: <http://www.unn.ac.uk/idr/report.html>.
78. Ainhoa Llorente Coto, Stefan Ruger, (2008) "Overview of Automated Image Annotation".
79. Duygulu, P., Barnard, K., de Freitas, J.F.G., Forsyth, DA, (2002) "Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary." In: *European Conference on Computer Vision*, London, UK, Springer-Verlag, 97-112.
80. Paul M. Mather, (2004). "Computer Processing of Remotely-Sensed Images: An Introduction.", John Wiley & Sons, England.
81. Fernández-Manso, A., Quintano, C., Santos, J.R., Maldonado, F, (2005). "Utilización de las imágenes CBERS-2 en el estudio de los sistemas forestales brasileños y su aplicabilidad en España.", *Actas del IV Congreso Forestal Español*, Zaragoza, 26-30 Septiembre 2005. Ed: Sociedad Española de Ciencias Forestales. ISBN: 84-921265-7- 4. Depósito legal Z-2577-05.
82. G. Viegas, V. Artola, C. Discoli, G. San Juan, (2008). "Exploración de herramientas para la interpretación de objetos a partir de imágenes aplicados al relevamiento de mosaicos urbanos.", *Avances en Energías Renovables y Medio Ambiente Vol. 12*.
83. Walker, E and K Okuma, (2000). "Automatic extraction of invariant features for objects recognition.", In Proceedings of the Annual Conference of the North American Fuzzy Information Processing Society, Boston, pp. 163–172. GA.
84. Jean-Baptiste Bordes, Michel Roux, (2006). "Detection of roundabouts in satellite image.", in *ISPRS, Ankara (Turkey)*.
85. Jorge Gabriel, Hoyos Pineda, (2007). "Aplicación de algoritmos genéticos a la clasificación de imágenes de satélite en el marco de los servicios Grid Inteligente Estado del Arte.", *Ciencia e Ingeniería Neogranadina*, Vol. 17-2, pp. 95-109. Bogotá. ISSN 0124-8170.

86. Sheikholeslami, G., et al., (1975). "Geographical Image Classification and Retrieval.", P. o. t. t. I. W. o. A. i. G. I. Systems. Las Vegas, USA. pp. 58-61.
87. Chuvieco, E. (2002). "Teledetección Ambiental.", Barcelona, Ariel Ciencia.
88. Tso, B. y Mather, P. (2001). "Classification Methods for Remotely Sensed Data.", New York, Taylor & Francis Inc.
89. Grossman, D. y Domingos, P., (2004). "Learning Bayesian Network Classifiers by Maximizing Conditional Likelihood.", En Proceeding of de 21th International Conference on Machine Learning. Banff, Canada.
90. Cetin, M., et al., (2004). Classification of Multi-Spectral, Multi-Temporal and Multi-Sensor Images Using Principal Components Analysis and Artificial Neural Networks: Beykoz Case. P. o. X. I. Congress. Estambul, Turkia. pp. 951-956.
91. Rosin, P. L., (2006). "Training Cellular Automata for Image Processing.", En IEEE Transactions on Image Proccesing. Vol. 15, No. 7 pp. 2076-2086.
92. Quirin, A., et al., (2004). "Learning Classifier Systems for Hyperspectral Images Processing.", En IlliGAL Report NO. 2004023.
93. Larch, D., (1994). "Generic Algorithms for terrain categorization of Landsat images.", En *SPIE*. Vol. 2103, pp. 14-19.
94. Lee, C. Y. y Antonsson, E. K., (2000). "Dynamic Partitional Clustering Using Evolution Strategies.", En Proceedings of de Third Asia Pacific Conference on Simulated Evolution and Learning.
95. Lozano, J. A. y Larrañaga, P., (1998). "Aplicación de los algoritmos genéticos al problema del clustering jerárquico.", En *Inteligencia Artificial*, No. 5.
96. Vatsavai, R. R., et al., (2006). *Miner: A Suit of Classifiers for Spatial, Temporal, Ancillary, and Remote Sensing Data Mining. S. N. D. M. Workshop:, I. a. i. E. S. w. t. t.S. o. t. Inter- and S. a. A. face of Computing Science.
97. Rafael Oliva Santos, (2009). "Anotaciones Semánticas de Datos Geográficos."
98. Eva Marie Klien, (2008). "Semantic Annotation of Geographic Information."
99. C. Angueira, J. López, G. Barraza, S. Prieto, G. Angella, D. Prieto, (2008). "Estimación de cultivos mediante teledetección y SIG en el área de riego del Río Dulce.", Santiago del Estero, Argentina.
100. A. Degioanni, J. Cisneros, S. Rang, (2001). "Teledetección y SIG para la gestión hidrológica del territorio.", *Revista de Teledetección*.
101. Santiago Castaño Fernández, (1999). "Aplicaciones de la Teledetección y SIG al control y cuantificación de las extracciones de agua subterránea."
102. L.G García-Montero, D. Bravo, C. Pascual, J. García Cañete, A. García-Abril, J. Urbano, (2002). "Análisis del medio con teledetección y SIG y participación social: primera fase de una planificación territorial en el Estado de Miranda (Venezuela).", *Revista Forestal Española*.
103. Bruno Basso Tim R. McVicar, Byong Lee, Hojjat Yazdanpanah and H.P. Das, <http://www.wmo.int/pages/prog/wcp/agm/gamp/documents/chap12-draft.pdf>, (Julio 2009). "Remote sensing and GIS applications in agrometeorology".
104. Alonso, Renato http://www.mappinginteractivo.com/plantilla.asp?id_articulo=1485, , (2008). "Ontologías geoespaciales en el dominio forestal".
105. Frederico T. Fonseca, Max J. Egenhofer, Peggy Agouris, (2002). "Using Ontologies for Integrated Geographic Information Systems".
106. Rafael Oliva-Santos, Leonid Mendoza-Robaina, Cynthia Costales-Llerandi, Eduardo Garea-Llano, Francisco Maciá Pérez, (2009). "Modelo de Anotación Semántica para Sistemas de Información Geográfica", VI Congreso Internacional de Geomática.
107. Raubal, M. (2005). "Mappings for Cognitive Semantic Interoperability", Toppen and M. Painho (Eds.), *AGILE 2005 - 8th Conference on Geographic Information Science*, IGP, Lisboa, Portugal, pp 291-296.
108. Klien, E., Lutz, M., (2005). "The Role of Spatial Relations in Automating the Semantic Annotation of Geodata.", A. Cohn, D. Mark (Eds), *LNCC 3693*, pp 133-148.

RT_033, octubre 2010

Aprobado por el Consejo Científico CENATAV

Derechos Reservados © CENATAV 2010

Editor: Lic. Lucía González Bayona

Diseño de Portada: Di. Alejandro Pérez Abraham

RNPS No. 2142

ISSN 2072-6287

Indicaciones para los Autores:

Seguir la plantilla que aparece en www.cenatav.co.cu

C E N A T A V

7ma. No. 21812 e/218 y 222, Rpto. Siboney, Playa;

Ciudad de La Habana. Cuba. C.P. 12200

Impreso en Cuba

