

CLASIFICACIÓN

DIGITAL

DE

IMÁGENES

POR

SATÉLITE



3º ITT.SE
Sistemas de Telemédica

Álvaro Yébenes Gómez
Álvaro Giner Sotos

ÍNDICE:

1. Introducción teórica a la clasificación digital.

1.1 – Fase de Entrenamiento.

- 1.1.1 – Método supervisado.
- 1.1.2 – Método no supervisado.

1.2 – Fase de Asignación.

- 1.2.1 – Clasificador de mínima distancia.
- 1.2.2 – Clasificador de paralelepípedos.
- 1.2.3 – Clasificador de máxima probabilidad.
- 1.2.4 – Otros criterios de clasificación.

2. Realización de la práctica.

2.1 – Introducción.

2.2 – Fase de entrenamiento: Programa STAD.EXE

2.3 – Fase de clasificación: algoritmo de mínima distancia.

2.4 – Conclusiones.

1.- INTRODUCCIÓN TEÓRICA A LA CLASIFICACIÓN DIGITAL

La clasificación de una imagen digital consiste en categorizar una imagen multibanda. Se pasa de tener unos ND continuos medidos por los sensores a una escala nominal o categórica de manera que cada píxel pertenece a una categoría definida previamente. El ND de un píxel clasificado es, en definitiva, el identificador de la clase o categoría en la que se haya incluido. Estas clases pueden describir distintos tipos de cubiertas (variable nominal o categórica) o bien intervalos de una misma categoría de interés (variable ordinal). Un ejemplo del primer tipo sería una clasificación de distintos tipos de cubierta vegetal, en cambio uno del segundo tipo sería un intento de clasificar distintos niveles de daño producidos por un incendio.

Los primeros intentos de clasificación digital se basaban exclusivamente en los ND de la imagen que son una traducción digital de un flujo energético recibido por un sensor para una determinada banda del espectro. Este tipo de clasificación tiene como inconveniente que no siempre permite separar cubiertas, ya que pueden existir distintas categorías temáticas que tengan un comportamiento espectral similar. En estos casos es conveniente.

La clasificación digital de imágenes es , en cierta medida , parecida a la foto-interpretación. En este método, se identifica el patrón visual asociado a cada cubierta (categoría) de acuerdo a una serie de parámetros como son el tono, la textura, la forma, el contexto, la disposición, etc. y después se identifican sobre las fotografías las superficies correspondientes a estas categorías, mediante su semejanza con el patrón-tipo previamente identificado. Por último es necesaria una verificación de los resultados.

En las técnicas digitales de clasificación de imágenes se dan estos mismos pasos. De esta forma se distingue entre las siguientes fases:

1. Definición digital de las categorías (fase de entrenamiento).
2. Agrupación de los píxeles de la imagen en una de las categorías previamente definidas (fase de asignación).
3. Comprobación y verificación de los resultados.

1.1 – Fase de entrenamiento.

La clasificación digital comienza con la definición de las categorías que se pretenden distinguir en la imagen. Se trata de una clasificación basada en los valores numéricos. Por lo tanto se trata de obtener el rango de ND que identifica a cada categoría para todas las bandas que intervienen en la clasificación.

Las distintas categorías no se definen solo por un ND sino por un conjunto de ND próximos entre sí. Existe una cierta dispersión en torno al ND medio de cada categoría. Por esto la fase de entrenamiento trata de definir con rigor cada una de las categorías que se pretenden distinguir teniendo en cuenta su dispersión en la zona de estudio.

Esto se consigue seleccionando una muestra de píxeles en la imagen que representen adecuadamente a cada categoría. A partir de estos se determina el ND medio de cada clase y el rango en el que varían, para todas las bandas que intervienen en la clasificación. Las estimaciones posteriores se basan sobre la muestra seleccionada por lo cual conviene seleccionarla adecuadamente ya que los resultados de la clasificación están mucho más influidos por la definición previa de las categorías, que por el criterio por el cual éstas son diferenciadas en la imagen.

Los métodos de clasificación se pueden distinguir en dos grupos: supervisado y no supervisado. El método supervisado parte de un conocimiento previo del terreno del cual se seleccionan las muestras para cada una de las diferentes categorías. En cambio, en el método no supervisado se buscan automáticamente grupos de valores espectrales homogéneos en la imagen para que después el usuario intente encontrar las correspondencias entre esos grupos seleccionados automáticamente y las categorías que le puedan ser de interés.

A propósito de esto, es necesario distinguir entre los dos tipos de clases que pueden intervenir en la clasificación temática: informacionales y espectrales. Las clases informacionales se establecen por el usuario y forman la leyenda de trabajo que se intenta discriminar. Las segundas corresponden a grupos de valores espectrales homogéneos

(reflectividad similar) y se deducen de los ND de la imagen. Idealmente se debería producir una correspondencia perfecta entre una sola clase informacional, es decir, una sola clase de cobertura y una sola clase espectral pero esto es muy poco corriente y lo más normal es que se produzca una de las siguientes situaciones:

- I. Puede ocurrir que una categoría informacional este expresada en varias categorías espectrales. Por ejemplo, un bosque de pinos podría aparecer en dos clases espectrales: uno correspondiente a solana y otro a umbría. Para solucionar este problema hay que definir correctamente la dispersión espectral de cada clase perfeccionando el muestreo.

- II. Otro caso sería que dos o más clases informacionales compartan una sola clase espectral. Se podría solucionar fundiendo las dos categorías en una más general. Por ejemplo se podría elegir una clase denominada coníferas en lugar de distinguir entre distintos tipos de pinar, o urbana, frente a distintas densidades de edificación. En caso de no poder fundir las dos categorías en una sola debido a su heterogeneidad se podría analizar una imagen de otra fecha del año en la cual la confusión sea evitable.

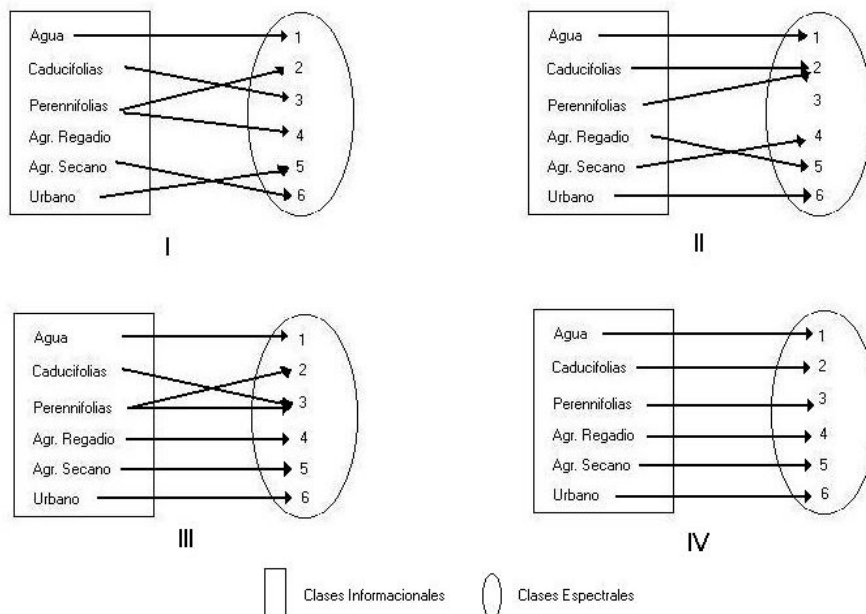


Fig. 1 Relaciones entre clases informacionales y espectrales: I una informacional-varias espectrales; II una espectral-varias informacionales; III mezcla de ambas; IV relación biunívoca entre ellas.

- III. También puede ocurrir que varias clases informacionales compartan varias clases espectrales. En este caso se debe replantear la estrategia de trabajo o probar alguna de las soluciones comentadas anteriormente.

Una vez expuesto esto, podemos afirmar que el método de entrenamiento supervisado pretende distinguir clases informacionales, mientras que el no supervisado trata de diferenciar las clases espectrales. Ninguno de los dos métodos es perfecto ya que en el método supervisado se puede forzar al ordenador a diferenciar categorías sin un claro significado espectral; y en el método no supervisado se pueden distinguir clases que no contengan una información de interés.

1.1.1 – Método supervisado.

Este método de entrenamiento requiere un cierto conocimiento de la zona de estudio que permite delimitar sobre la imagen unas zonas o áreas representativas de las distintas categorías que se pretenden discriminar. Estas áreas se conocen como *training fields* (áreas de entrenamiento) ya que sirven para entrenar al ordenador para que pueda reconocer las distintas categorías. A partir de estas áreas el ordenador calcula diversos parámetros estadísticos de los ND que definen cada clase, para luego clasificar el resto de píxeles en una categoría determinada atendiendo a sus ND. Resulta adecuado seleccionar varias áreas de entrenamiento por clase para reflejar correctamente su variabilidad en la zona de estudio.

Una vez acabada la selección de las áreas de entrenamiento, el ordenador debe proceder a calcular las estadísticas elementales de cada categoría: media, rango, desviación típica, matriz de varianza – covarianza, etc., a partir de los ND de todos los píxeles incluidos en el área de entrenamiento de dicha clase. Este cálculo es aplicado a cada banda espectral. Las medidas extraídas a partir de los ND de cada área de entrenamiento definen a cada clase, por lo tanto es muy importante su correcta selección ya que de otro modo se forzaría al ordenador a clasificar áreas heterogéneas.

Un sencillo ejemplo ilustra lo descrito anteriormente. Disponemos de una imagen en la que se desean distinguir 3 clases: viñedo, campos de cereales y encinar. La primera opción a la hora de definir áreas de entrenamiento sería seleccionar una por cada categoría. A partir de estas áreas de entrenamiento, el ordenador calcularía una serie de parámetros estadísticos que serían los que definirían a esa determinada clase.

La selección de áreas de entrenamientos sería incorrecta si alguna de estas se situara entre dos o más categorías ya que al tratarse de un área heterogénea, sus medidas estadísticas serían una cierta media entre las distintas clases. Otro error que se podría cometer sería seleccionar únicamente como áreas de entrenamiento aquellas zonas de la cubierta donde esta sea muy homogénea ya que entonces estaríamos dejando pasar la variación real que tenga dicha cubierta. Por ejemplo, al clasificar el encinar deberíamos seleccionar como áreas de entrenamiento tanto zonas donde el encinar aparezca con una densidad mayor como zonas donde la densidad sea menor.

Respecto al tamaño que deben tener las áreas de entrenamiento, se requiere que tengan un tamaño de $m + 1$ píxeles por categoría, siendo m el número de bandas espectrales que intervienen en la clasificación. Es conveniente de todos modos superar ampliamente este mínimo. Además, es mejor elegir varias áreas de pequeño tamaño que una sola de mayores dimensiones. Otros métodos que se pueden utilizar son seleccionar píxeles aislados de cada categoría o señalar píxeles puros en torno a los cuales el ordenador crea una parcela de modo automático en función de la distancia o la similitud espectral (o ambas) de los píxeles vecinos del inicialmente seleccionado.

1.1.2 – Método no supervisado.

Este método de clasificación trata de definir las clases espectrales presentes en la Imagen. No implica ningún conocimiento previo del área de estudio por lo que la intervención humana se centra en la interpretación de los resultados.

Este método asume que los ND de la imagen se agrupan en una serie de conglomerados (o clusters) que se corresponden con grupos de píxeles con un

comportamiento espectral homogéneo y que, por ello, deberían definir unas clases informacionales de interés. Por desgracia estas clases espectrales no pueden ser asimiladas siempre a las categorías temáticas que el usuario pretende deducir por lo que es labor de éste interpretar el significado temático de dichas categorías espectrales.

El método para definir los grupos con un comportamiento espectral similar es muy parecido a otras técnicas de clasificación automática de datos. Se basa en la selección de tres parámetros: variables que intervienen en el análisis, criterio para medir la distancia o similitud entre casos, y un criterio para agrupar los casos similares. En nuestro caso las variables serían las diferentes bandas espectrales. Los casos son los píxeles que componen la imagen y cada uno de ellos está definido por tantos ND como bandas tengamos. Lo que trata este método es encontrar grupos de píxeles con ND parecidos para asignarlos a alguna de las categorías temáticas que deseemos.

La delimitación de cada grupo espectral se inicia señalando dos criterios: uno que marque la similitud entre píxeles, y otro que marque las condiciones del proceso de agrupamiento. Para el primero, el método más utilizado se basa en la distancia euclídea aunque se pueden utilizar otros criterios como el de la distancia media o la de Mahalanobis. En cuanto al algoritmo de agrupamiento, el más extendido es el denominado ISODATA.

1.2 – Fase de asignación.

En este momento vamos a exponer los principales algoritmos de clasificación. Desde un punto de vista estadístico, todos los algoritmos definen un área de dominio de cada clase en torno a su centro mediante un conjunto de funciones. Un determinado píxel será asignado a una clase si sus ND se encuentran dentro de los límites establecidos para dicha clase.

1.2.1 - Clasificador de mínima distancia.

El criterio más sencillo para clasificar un píxel en una categoría es incluirlo en la más cercana a él, es decir, en aquella que minimice la distancia entre ese píxel y el centroide de clase. Esta distancia no es una distancia geográfica sino espectral consecuencia de comparar los ND de cada píxel con los del centro de las distintas categorías, para todas las bandas que intervienen en el análisis.

Existen varias formas de medir esta distancia espectral entre píxeles y clases pero la más empleada es la distancia euclídea:

$$d_{ij}^{(k)} = \sqrt{\sum_{n=1}^{N_{bandas}} (X_{ij}^n - \bar{X}_{ij}^{(k)})^2} \quad (\text{distancia Euclídea})$$

donde $k=1,2,3,\dots,N_{clases}$

Siendo X_{ij}^n ($n=1,2,3,\dots,N_{bandas}$) el valor de cada píxel en la banda n , y $\bar{X}_{ij}^{(k)}$ el valor del "centroide" (media aritmética) de la clase k en la banda n . Los subíndices (i,j) corresponden al número de fila y número de columna en cada píxel de la imagen.

Una vez determinadas las distancias a cada clase, se toma la mínima de todas ellas, y se asigna el píxel correspondiente a la clase más próxima (es decir, a la clase para la cual la distancia es menor).

$$clase(i, j) = \{k \mid d_{ij}^{(k)} = \text{mínimo}\}$$

Este algoritmo es relativamente rápido de ejecutar y ofrece buenos resultados cuando no exista un gran solape entre clases. Además, no deja ningún píxel sin clasificar puesto que siempre existe una clase más cercana.

1.2.2 – Clasificador de paralelepípedos.

En este método, el usuario fija un área de dominio para cada categoría teniendo en cuenta sus valores de centralidad y dispersión. Posteriormente cada píxel será asignado a una clase si sus ND están dentro del área de dominio de esa clase, para todas las bandas que intervienen en la clasificación. El píxel x será asignado a la clase A si sus ND en las distintas bandas ($ND_{x,k}$) están incluidos en el área de dominio de esa clase:

$$\overline{ND}_{A,n} - R_{A,n} \leq ND_{x,n} \leq \overline{ND}_{A,n} + R_{A,n}$$

para todo $n=1,2,3,\dots,N$ bandas

$R_{A,n}$ = rango de dispersión señalado para la categoría A en cada banda n.

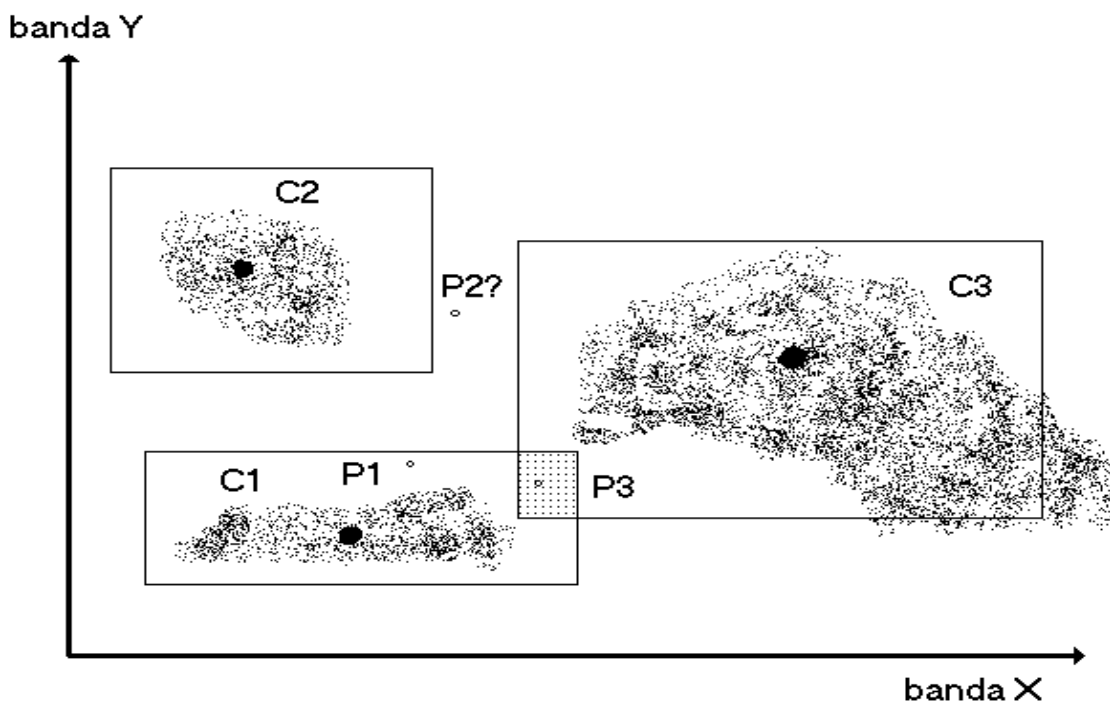


Fig. 2 Clasificador de paralelepípedos. En esta figura se observan algunos de los problemas de este algoritmo: el píxel P1 queda clasificado sin problemas en la clase C1; sin embargo el píxel P2 no está encuadrado en ningún área de dominio por lo que quedará como no clasificado. El píxel P3 se encuentra en una zona de solape entre el dominio de dos categorías por lo que podrá asignarse aleatoriamente a dos categorías distintas.

Este algoritmo es muy rápido de ejecutar ya que solamente necesita operadores condicionales del tipo IF... THEN. Pero presenta ciertos problemas inherentes al diseño de las áreas de dominio ya que pueden existir ciertos píxeles que queden sin clasificar y otros que se encuentren en zonas comunes a dos o más categorías. En este último caso el ordenador asignará este píxel a la clase que se encuentre antes en el algoritmo de clasificación, aunque se puede utilizar conjuntamente el algoritmo de mínima distancia para evitar el solape. La solución para los píxeles sin clasificar es ampliar el rango de dispersión y repasar la definición espectral de las categorías ya que la presencia de píxeles no clasificados indica que las clases no están bien definidas o no incluidas en la leyenda de trabajo.

1.2.3 – Clasificador de máxima probabilidad.

Este método considera que los ND de cada categoría se ajustan a una distribución normal. Esto nos permite describir esa categoría por una función de probabilidad, a partir de su vector de medias y matriz de varianza-covarianza. Así podemos calcular la probabilidad de que un determinado píxel pertenezca a una categoría; el cálculo se realiza para todas las categorías definidas y el píxel se clasifica en clase a la cual tenga más probabilidad de pertenecer.

Este clasificador es el más complejo y, por lo tanto, el que mayor volumen de cálculo requiere, pero es el más robusto y fiable ya que se ajusta fielmente a la distribución original de los datos.

Para calcular la probabilidad asociada a un cierto ND se necesita la media y la desviación típica de cada categoría. En el caso de trabajar con una sola banda la probabilidad se calcula de la siguiente forma:

$$p(x/A) = \frac{1}{\sqrt{2\pi\sigma_A^2}} e^{\left\{-\frac{(ND_x - \overline{ND}_A)^2}{2\sigma_A^2}\right\}}$$

$p(x/A)$ = probabilidad de que un cierto píxel x (definido por ND_x) pertenezca a la clase A ,

\overline{ND}_A = media de la clase A

σ_A^2 = varianza de la clase A

Una vez determinada la probabilidad para todas las categorías el píxel será asignado a la clase A, sólo si:

$$p(x/A) \geq p(x/B)$$

$\forall A \neq B$, con $B = 1, 2, 3, \dots, m$; m = número de categorías

Para el caso de dos bandas, los límites de las categorías tienen una forma elipsoidal (Fig. 3). Cada una de estas elipses es una línea de isoprobabilidad, es decir todos sus puntos tienen la misma probabilidad de pertenecer a la categoría asociada.

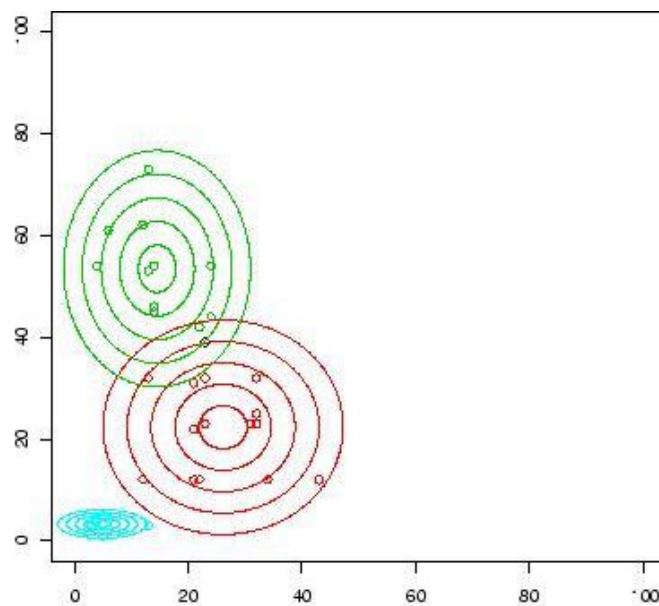


Fig. 3 : Extensión a dos bandas del algoritmo de máxima probabilidad

Con más de dos bandas resulta imposible representar gráficamente este criterio de clasificación y se debe extender la fórmula de la probabilidad. Una vez que el ordenador haya asignado los píxeles puede establecerse algún criterio que permita asignarlos a una clase sólo si la probabilidad correspondiente es superior a determinado umbral. Este criterio asume que las categorías presentan una distribución normal; sin embargo, la hipótesis de que los datos de reflectividad siguen una distribución normal no siempre se cumple y debería verificarse siempre.

1.2.4 – Otros criterios de clasificación.

A continuación se exponen brevemente otros clasificadores:

- Clasificador en árbol: suponen un análisis exhaustivo de las respuestas espectrales de las clases y del conjunto de datos disponibles, consiste en ir haciendo preguntas a cada píxel cuya respuesta positiva o negativa conducirá a otra pregunta y así sucesivamente hasta obtener la clase a la que pertenece. Se basa en los mismos principios que los sistemas expertos.
- Clasificador de contexto espacial: todos los métodos estudiados hasta ahora utilizan, para clasificar un píxel, sólo los valores de reflectividad recogidos en dicho píxel. Este hecho no es demasiado problemático si el tamaño de píxel es mayor que los elementos que los diferentes elementos físicos que componen el paisaje. Sin embargo si el tamaño del píxel es más pequeño que las unidades de paisaje podría utilizarse la información de los píxels de alrededor para estimar la pertenencia a una clase o confirmar la estimación. Puede resultar un método útil en combinación con el de máxima probabilidad ya que permite incluir la información de los píxels circundantes para tomar una decisión en caso de que las probabilidades para dos clases sean similares o no exista una clase con probabilidades de pertenencia suficientemente altas.

Se trata de incorporar otras fuentes de información distintas a las bandas para la clasificación. Entre esta información estaría la altitud, pendiente, litología, etc. Las distintas formaciones vegetales tienen mayor capacidad para desarrollarse en determinados entornos definidos en parte por estas variables. Por tanto puede ser útil establecer cual es la formación vegetal con mayor probabilidad de desarrollarse en un determinado píxel (en función de topografía y litología) y combinar esta información con la de las bandas.

- Clasificación borrosa: los métodos anteriormente expuestos consideran que un píxel solamente puede pertenecer a una determinada clase. Por ejemplo, en el clasificador de máxima probabilidad, un píxel es clasificado en la clase más probable; da lo mismo que la pertenencia a esta clase sea mucho mayor que al

resto (90% frente a un 10%) o que la diferencia sea menor (55 frente a 45). En el primer caso se estaría cometiendo un error muy bajo al clasificar el píxel, en cambio en el segundo la asignación sería muy arriesgada. En estos casos tiene más sentido no clasificar los píxels de forma unívoca sino establecer cual es su *posibilidad* de pertenencia a cada una de las clases (el concepto de posibilidad no es exactamente igual al de probabilidad). El píxel ahora no pertenece a una sola clase totalmente sino que tiene distintos grados de pertenencia comprendidos entre 0 y 1, lo que permite una asignación simultanea a varias categorías.

- Clasificadores basados en redes neuronales: se basan en el uso de redes neuronales artificiales que, se supone, imitan a las redes neuronales reales en el desarrollo de tareas de aprendizaje. Una neurona artificial es un objeto lógico (se trata de software no de hardware) que recibe diversas entradas, hace una suma ponderada de las mismas y produce una salida a partir de la aplicación de una función umbral a la media ponderada.

Si conectamos las salidas de unas neuronas como entradas de otras obtenemos una red neuronal. Uno de los ejemplos más típicos de red neuronal es la *Back Propagation Neural Network* que aparece en el siguiente gráfico.

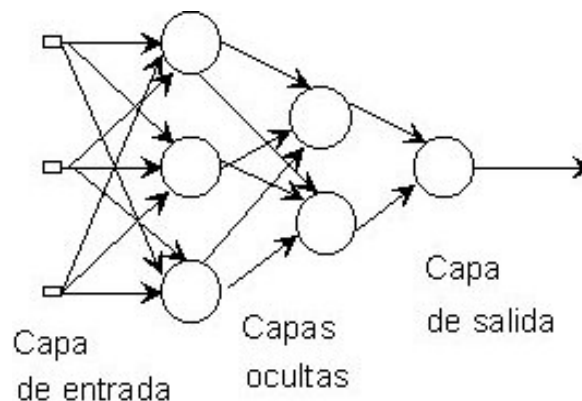


Fig. 4: Red neuronal BPN

Consta de una capa de entrada con tantas neuronas como variables de entrada se vayan a introducir en el modelo (en el caso de la teledetección sería una por cada banda utilizada para clasificar), una capa oculta que realiza la mayor parte del cálculo y una capa de salida con tantas neuronas como posibles clases existan. En teledetección esta salida suele consistir en un valor numérico entre 0 y 1 para cada

clase, cuanto mayor sea este valor más verosímil resulta que el pixel pertenezca a la clase en cuestión

Para trabajar con una red neuronal existen varias fases:

- **Entrenamiento.** Se le introducen a la red la respuesta espectral de pixeles cuya clase se conoce y se compara la salida con la realidad. A partir de esta comparación se modifican los coeficientes de ponderación de todas las neuronas para que se obtenga la respuesta adecuada (se trata de un procedimiento automático) es decir un 1 en la clase correcta y ceros en las incorrectas
- **Estabilización.** Al principio del entrenamiento, los factores de ponderación cambian muy deprisa, pero conforme este se desarrolla (y si las areas de entrenamiento se han seleccionado correctamente) se estabilizan (no se modifican aunque se vuelvan a introducir los pixels de entrenamiento). En este momento finaliza la fase de entrenamiento
- **Clasificación** Se introducen las respuestas espectrales de los pixels cuya clase no se conoce y se adjudican a la clase que de una respuesta más alta (que no va a ser necesariamente 1).

Se trata en definitiva de un método de clasificación robusto que da buenos resultados cuando las respuestas espectrales de las clases no siguen una distribución normal. La clave está en el conjunto de coeficientes de ponderación que constituyen un conjunto de parámetros que deben ajustarse a unos datos de entrada y salida. Por tanto en cierto modo es equivalente a una regresión.