

## Integración de Técnicas de IA y Teledetección para la localización de pesquerías.

A. Iglesias<sup>(1)</sup>, J.M. Cotos<sup>(2)</sup> y B. Arcay<sup>(1)</sup>

<sup>(1)</sup>Dep. Tecnologías de Información y Comunicaciones. Universidad de La Coruña. La Coruña. Campus de Elviña s/n, 15071, La Coruña. elfonso@udc.es, cibarcay@udc.es.

<sup>(2)</sup>Dep. Electrónica y Computación. Universidad de Santiago Compostela. Instituto de Investigaciones Tecnológicas. Campus Sur. Santiago Compostela.. manel.cotos@usc.es.

### Resumen

En este documento se recogen las bases de un sistema de apoyo a la toma de decisiones sobre localización de pesquerías basado en técnicas de IA. El propósito de este sistema es el ahorro de costes a las flotas pesqueras sin dañar el desarrollo sostenible de los recursos naturales. Para ello, emplearemos como fuentes de datos por un lado imágenes de concentración de clorofila (OrbView-2), temperatura superficial (Serie NOAA) y anomalía altimétrica (Topex/Poseidon), y por otro datos reales de capturas obtenidas del cuaderno de pesca de un barco piloto (palangrero del Atlántico Norte). Hemos realizado finalmente una comparativa de técnicas de IA y exportado los resultados a un SIG.

### 1. Introducción

La pesca, entendida como la explotación de los recursos marinos, es una de las principales actividades económicas en España en general y en Galicia en particular. Sin embargo, este importante sector económico se ha visto amenazado en estos últimos años por desastres ecológicos, políticas de cuotas y paros biológicos. Un sector tan competitivo como el de la pesca debería de incorporar la aplicación de nuevas tecnologías como sistemas de apoyo a la toma de decisión, siempre dentro de un marco de desarrollo sostenible [1].

En este ámbito la teledetección es una fuente de información vital. Proponemos relacionar la información procedente de imágenes de OrbView-2, serie NOAA y Topex-Poseidon con datos de capturas a través de técnicas de Inteligencia Artificial. Para ello contamos con datos reales de capturas proporcionados por un barco colaborador que se dedica a la pesca con palangre en el Atlántico Norte. Con ellos y la información procedente de los sensores remotos estamos en disposición de crear un conjunto de entrenamiento y otro de validación para comparar los resultados de diferentes predictores basados de forma genérica en sistemas conexionistas. La consecución de parámetros válidos para nuestros predictores permitirá sugerir al patrón de pesca zonas más probables que conducirían a la reducción los tiempos de búsqueda de

pesquerías, con el consiguiente ahorro de tiempo y combustible que ello conlleva, aumentando dentro de las posibilidades legales el volumen de capturas.

### 2. Métodos y fuentes de datos

En este apartado explicamos los fundamentos de un nuevo sistema de apoyo a pesquerías. Concretamente, a partir de los datos obtenidos de varios sensores remotos y del cuaderno de pesca de un barco colaborador, nuestro sistema (ver [2]) proporciona la probabilidad de conseguir capturas de una especie de tiburón pelágico, el *Prionace Glauca*, más conocido como tiburón azul o quenlla. El sistema desarrollado es extensible a cualquier especie marina de la cual se posea suficiente información para el entrenamiento del sistema.

Este sistema se puede descomponer en tres fases diferenciadas, como se puede ver en la figura 1, cuyo objetivo final es transmitir productos susceptibles de ser enviados a las unidades embarcadas.

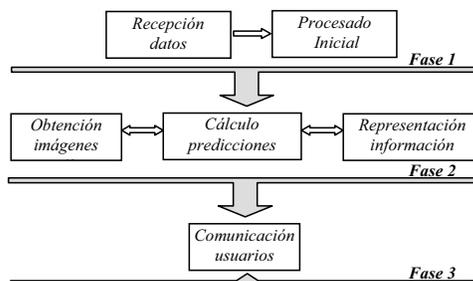


Figura 1.- Organigrama del Sistema de Apoyo a pesquerías

La fase 1 tiene como misión principal la recepción de información procedente de varios satélites. El sistema se compone de un parque de antenas que permiten la recepción de distintos satélites como NOAA (en alta y baja resolución), OrbView-2 (imágenes del sensor SeaWifs) y Meteosat. Los

datos recogidos por las antenas se almacenan en un sistema de *backup* y se distribuyen a todos los equipos que están formando parte de una red local. Cualquiera de estos equipos tiene acceso a los datos para poder visualizarlos y procesarlos. Puede encontrarse una descripción exhaustiva de esta fase en [1], [2], [3]

El objetivo de este trabajo es, empleando los datos recibidos como resultado de la fase 1, elaborar la *fase 2*, que consistirá en:

- Aplicar técnicas de procesamiento digital sobre los productos iniciales para conseguir nuevos datos con significado biológico, por ejemplo los siguientes filtros pasa-alta para detectar la existencia de frentes térmicos:
  - DoG (ver [4]).
  - Cluster-Shade (ver [5]).
- Visualizar esta información independientemente de la plataforma usada.
- Estudio de la sensibilidad y correlación de los datos iniciales empleando:
  - Análisis de componentes principales.
  - Mapas Autoorganizativos de Kohonen.
- Calcular la probabilidad de obtener capturas pesqueras en función de parámetros medioambientales obtenidos mediante teledetección. Para ello probaremos las siguientes técnicas:
  - Redes entrenadas con el algoritmo de retropropagación (*backpropagation*).
  - Redes RBF (Radial Basis Function).
  - Redes funcionales.
  - ANFIS (Sistema de inferencia neuro-difuso)
- Gestionar toda la información de manera eficiente y con un control centralizado. Para ello, diseñaremos una base de datos apropiada siguiendo el modelo E-R.

En la *fase 3*, nuestro sistema se comunica con los usuarios. La información puesta a disposición de la unidad embarcada puede ser enviada haciendo uso de diversos medios de comunicación, por ejemplo la red global por satélite Inmarsat. Esta red de satélites proporciona una gran diversidad de servicios: telefonía, datos, fax, web,... Es usada también para comunicaciones de desastre o emergencia marítima, por lo que no resulta infrecuente hallarse con barcos que cuentan con este servicio de comunicaciones.

Además, la disponibilidad de conexión a la red Internet permite realizar el envío de información haciendo uso del correo electrónico. El software embarcado TUNAFIS 4.2 desarrollado por otros miembros del *Instituto de Investigaciones Tecnológicas* permite gestionar cuentas de usuario de correo electrónico y extraer la información enviada para su gestión en el interfaz gráfico.

### 2.1. Fuentes de datos

Hemos empleado para este trabajo imágenes de la serie NOAA, OrbView-2 y Topex-Poseidon. Después de un procesamiento inicial los productos a usar como entradas serán los siguientes:

- Temperatura superficial o SST (NOAA)
- Anomalía térmica (NOAA)
- Frente térmico (NOAA)
- Concentración superficial de clorofila (OrbView-2)
- Anomalía altimétrica (Topex-Poseidon)

Tras realizar un estudio de sensibilidad previo mediante redes de Kohonen, se llegó a la conclusión de que la anomalía térmica no aportaba información de relevancia, probablemente porque es redundante con la SST, por lo que la hemos suprimido del estudio.

Además de los datos conseguidos del procesamiento de imágenes de satélite, tenemos datos de campo procedentes de un barco pesquero colaborador. El método de pesca es el palangre y la información anotada consiste en las coordenadas geográficas del inicio y final del lance de palangre, así como las capturas de cada especie recogidas diariamente a lo largo de los años 1998 y 1999.

Por tanto, disponemos de un conjunto de entrenamiento formado por cuatro entradas procedentes de imágenes de sensores remotos y un *a salida*, que es el número de capturas de una determinada especie marina.

### 2.2. Métodos

La falta de modelos matemáticos y de una colección clara de reglas de conocimiento, así como la existencia de un conjunto de datos de entrenamiento apoya el uso de sistemas conexionistas dentro de las diferentes técnicas de Inteligencia Artificial.

#### 2.2.1. Redes Neuronales

Las redes neuronales han recorrido un largo camino desde las primeras publicaciones, pero actualmente se han estabilizado como una materia interdisciplinar, con profundas raíces en neurociencias, psicología, matemáticas, física e ingeniería.

Trabajos anteriores de otros autores avalan las redes neuronales como metodología adecuada para la obtención de resultados fiables. Citaremos algunos:

- *T. Komatsu et al* [6] realiza predicciones de capturas de sardina japonesa usando R. N. Con el análisis de los pesos sinápticos intentan analizar cuál es el factor físico o biológico de mayor importancia. Se usaron antes modelos de regresión sin resultados satisfactorios debidos a la correlación entre las variables de entrada. Con R. N., en cambio sí se han conseguido resultados satisfactorios.

- *D. Aurelle et al* [7] realiza predicciones de datos pesqueros empleando un perceptrón con 3 capas y 2 neuronas en la capa oculta, entrenado con el algoritmo de retropropagación de error.

- *M. J. Dreyfus-Leon* [8] predice el comportamiento de un pescador usando redes neuronales.

- *Aussem y Hill* [9] predicen la presencia de una dañina alga verde (*Caulerpa taxifolia*) mediante un perceptrón multicapa con aprendizaje supervisado.

- *Brosse et al* [10] realiza predicciones de abundancia de pesca en lagos usando redes neuronales. Se presenta una comparación de MLR (Regresión Lineal Múltiple) con RN y con un análisis de componentes principales, con la cual se demuestra que las redes neuronales predicen de manera más precisa.

- *Maas et al* [11] realiza predicciones de parámetros medio-ambientales. Para ello emplea series temporales correspondientes al fenómeno de El Niño.

Una Red de Neuronas Artificiales puede definirse, de forma general, como una máquina que es diseñada para modelar la forma en la que el cerebro humano realiza una tarea o función de interés; la red neuronal es normalmente implementada utilizando componentes electrónicos o simulada por software utilizando un ordenador. Para obtener estos resultados, las redes emplean interconexiones masivas entre elementos simples de procesamiento llamados neuronas.

### 3. Resultados

### 3.1. Elección del algoritmo

Partiendo de 4 variables de entrada (SST, Calentamiento-Enfriamiento, concentración de clorofila, altimetría) y una salida (capturas de Quenlla), creamos un perceptrón multicapa entrenado con algoritmo de retropropagación de error (Backpropagation) [12] [13]:

El objetivo es encontrar una red lo más sencilla posible, que nos permita extraer conclusiones sobre las condiciones en las cuales se puede maximizar el número de capturas de la unidad embarcada.

Una vez entrenada la red, se comprueba que el número de neuronas de la capa escondida apenas influye en el error que presenta la red. Existen pequeñas diferencias iniciales, que pueden ser debidas a la inicialización aleatorio de los pesos de la red. En todos los casos, el error MSE del conjunto de entrenamiento permanece cercano a 0.01, mientras que el error del conjunto de validación es ligeramente inferior a 0.02.

Ante estos resultados, nos decantaremos para continuar nuestro trabajo con la red más sencilla, es decir, aquella que posee 2 neuronas en su capa escondida. Además, hemos comparado el error cuadrático medio del conjunto de entrenamiento y validación de otros paradigmas modernos dentro de la IA, las redes funcionales [14] y el sistema neuro-difuso ANFIS [15]. Los resultados vienen resumidos en la figura 2.

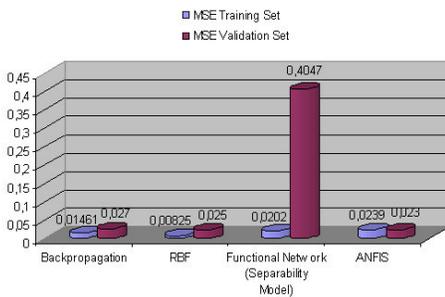


Figura 2.- MSE de los conjuntos de entrenamiento y validación en los 4 algoritmos probados.

Para que el algoritmo estudiado sea un buen predictor, hemos de conseguir un error bajo para el conjunto de entrenamiento y el de validación. Cuando se da el caso de que el error de validación es mucho mayor que el error de entrenamiento, estamos ante el fenómeno sobreajuste o sobreentrenamiento. Es un hecho bastante estudiado el que cuando se utiliza un modelo con muchos parámetros para ajustar un conjunto de datos procedente de un proceso con pocos grados de libertad, el modelo obtenido puede no descubrir las tendencias reales del proceso original, aunque pueda presentar un error pequeño de ajuste de datos. En este caso, el entrenamiento se reduce a la interpolación de los datos, incluyendo el ruido, utilizando una función sigmoideal complicada.

En la figura 2 hemos escogido los mejores resultados después de entrenar los diferentes casos de redes neuronales, funcionales y ANFIS. Para elegirlos, se han evitado todos aquellos casos en los cuales existía el problema de sobreajuste. Los errores representados en la figura 2 son todos del mismo orden exceptuando el error de validación de la red funcional. Ello quiere decir que la capacidad de generalización de este algoritmo para nuestro problema es

menor, probablemente debido a la no existencia de un modelo matemático capaz de explicar el problema [16].

Finalmente, hemos escogido para la implementación el perceptrón multicapa entrenado con algoritmo Backpropagation (retropropagación de error) por su sencillez frente al ANFIS y la red RBF.

### 3.2. Generación de mapas de probabilidad de pesca

Para generar los mapas de probabilidad, hemos de tratar con los diferentes productos de partida. El formato de datos empleado por la serie NOAA es el *tdf*, el de Orbview-2 el *hdf*, mientras que las imágenes de Topex-Poseidon están en binario. Los lenguajes de programación empleados para tratar las imágenes fueron C en el caso de imágenes NOAA, e IDL 5.2 para imágenes Orbview-2. Creemos conveniente unificar el formato de las imágenes como paso previo a la aplicación de la red neuronal, por ello hemos pasado las imágenes *hdf* y *binarias* a formato *tdf*, para luego tratarlas con código generado en lenguaje C. Hemos escogido el formato *tdf* por disponer de un conjunto versátil de funciones especialmente indicadas para tratamiento de imágenes en el citado formato. Es importante ajustar la corrección geográfica de las imágenes para que exista una correspondencia píxel a píxel entre ellas, y facilitar la aplicación del algoritmo correspondiente.

Para evitar problemas de este tipo, hemos corregido las imágenes NOAA y OrbView-2 mediante las correcciones rectangular y cilíndrica equidistante, respectivamente. Ambas correcciones geográficas son equivalentes píxel a píxel si poseen la misma resolución espacial. Para la imagen de Topex-Poseidon, hemos calculado computacionalmente las coordenadas geográficas de cada punto para colocarlas después sobre una imagen inicialmente sin datos.

Estamos en condiciones de aplicar la red Backpropagation para calcular el mapa de probabilidad de obtener capturas del Prionace Glauca. Las entradas de la red serán los valores de los píxeles de cada imagen. Para la entrada correspondiente a la anomalía altimétrica tomaremos como dato el valor no nulo más cercano al píxel, si es que no existe ya. Con la salida de la red generamos un mapa de probabilidad de pesca que estamos en condiciones de exportar a un SIG (ver figura 3). Con niveles de gris altos (colores claros) tenemos una probabilidad alta de pesca.

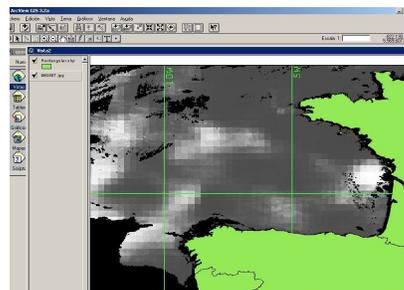


Figura 3.- Mapa de probabilidad de capturas de Prionace Glauca correspondiente al 7-8-1998 conseguido mediante una red neuronal entrenada con el algoritmo Backpropagation. Con niveles de gris altos (colores claros) tenemos una probabilidad alta de pesca.

- ✓ Se ha elaborado un sistema de apoyo a la toma de decisiones para la explotación operacional de pesquerías. Para ello, hemos integrado diferentes técnicas de Inteligencia Artificial dentro de un sistema de adquisición de datos procedentes de sensores remotos.
- ✓ A partir de las imágenes de dos años completos (98 y 99) y de datos de campo, se ha generado un amplio sistema de información que incluye una base de datos relacional formada por parámetros medioambientales, coordenadas geográficas y capturas.
- ✓ Empleando el sistema de información antes mencionado, hemos diseñado nuestro sistema de apoyo a la toma de decisiones. Nuestro problema se caracteriza por la falta de modelos globales y locales, la inexistencia de una base de conocimiento y por unas variables poco interrelacionadas (si excluimos las anomalías y frentes térmicos). Debido a estas razones, nos hemos decantado por algoritmos propios de la tendencia conexionista dentro del campo de la Inteligencia Artificial.
- ✓ Hemos conseguido diseñar, por tanto, varias herramientas capaces de predecir cuáles son las zonas óptimas de pesca en función de la información aportada por diferentes satélites. Los resultados obtenidos han sido validados empleando patrones distintos de los usados en el entrenamiento.
- ✓ Se han desarrollado las aplicaciones necesarias que integran las redes neuronales dentro del tratamiento digital de las imágenes. Para ello, ha sido necesario unificar previamente los distintos formatos de imágenes con los que trabaja cada satélite. La salida del sistema consiste en un mapa de probabilidad de pesca generado a partir de la salida de las red empleada. De forma simultánea se han implementado aplicaciones para el acceso al sistema de información y para el cálculo de predicciones puntuales. Estas herramientas han sido desarrolladas de forma que un usuario del sistema disponga de todas las utilidades en un entorno cómodo y accesible para él.

- [4].- Wells W. M. *Efficient Synthesis of Gaussian Filters by Cascaded Uniform Filters*. IEEE Trans. Pattern Analysis and Machine Intelligence-9 No. 2 March 1986.
- [5].- Harlow C. A., Trivedi M. M., y Connors R. W. *Use of texture operators in segmentation*. Optical Engineering, vol. 25 , no. 11, pp. 1200-1206, Nov. (1986)
- [6].- Komatsu T., Aoki I., Mitani I., y Ishii T. *Prediction o the Catch o Japanese Sardine Larvae in Sagami Bay Using a Neural Network*. Fisheries Science 60(4),385-391 (1994).
- [7].- Aurelle D., Lek S., Giraudel J., Berrebi P. *Microsatellites and artificial neural networks: tools for the discrimination between natural and hatchery brown trout (Salmo trutta, L.) in Atlantic populations*. Ecological Modelling 120 313-324. (1999)
- [8].- Dreyfus-Leon M. J. *Individual-based modelling of fishermen search behaviour with neural networks and reinforcement learning*. Ecological Modelling 120 287-297. (1999)
- [9].- Aussem A., Hill D.. *Neural-network metamodelling for the prediction of Caulerpa taxifolia development in the Mediterranean sea*. Neurocomputing 30 ; 71-78 (2000)
- [10].- Brosse S., Guegan J., Tourenq J., Lek S. *The use of artificial neural network to assess fish abundance and spacial occupancy in the litoral zone of a mesotrophic lake*. Ecological Modelling 120:299-311. (1999)
- [11].- Maas O., Boulanger J., Thiria S. *Use of neural networks for predictions using time series: Illustration with the El Niño Southern oscillation phenomenon*. Neurocomputing 30: 53-58. (2000)
- [12].- Gorman R. P. y Sejnowski, T. J. *Learned Classification of Sonar Targets Using a Massively Parallel Network*. IEEE Transactions on Acoustics, Speech, and Signal Processing. 36:1135-1140. (1998)
- [13].- Rumelhart D.E., Hinton, G. E. y Williams, R. J. *Learning internal representations by errors propagation. In Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. 1. D.E. Rumelhart and J.L. Mac Clelland, Cap. 8. MIT Press. (1986)
- [14].- Castillo E. and Gutiérrez J.M. *Nonlinear Time Series Modeling and Prediction Using Functional Networks. Extracting Information Masked by Chaos*. Physics Letters A, Vol. 244, 71-84 (1998).
- [15].- Jang J.-S.R. *ANFIS: adaptive-network-based fuzzy inference system*. IEEE Trans. Syst., Man., Cybern, vol.23, no.5, pp. 665-685. (1993)
- [16].- Iglesias A., B. Arcay, J.M. Cotos. *Optimisation of fishing predictions by means of Artificial Neural Networks, ANFIS, Functional Networks and Remote Sensing images*. Expert Systems with Applications. Aceptado y pendiente de publicación.

## 5. Referencias

- [1].- Triñanes J. A. *Sistema de información basado en teledetección para ayuda a la explotación operacional de pesquerías de túnidos y otras especies pelágicas*. Tesis doctoral. Departamento Electrónica e Computación, Universidad de Santiago. (Febrero 1998)
- [2].- Iglesias A. *Sistema de apoyo a la explotación operacional de pesquerías basado en técnicas de inteligencia artificial y teledetección*. Departamento de Electrónica y Computación, Facultad de Física, Universidad de Santiago de Compostela. (Mayo 2003)
- [3].- Cotos J. M. *Dinámica y clasificación de estructuras oceánicas para aplicación operacional de pesquerías utilizando teledetección e ingeniería de conocimiento*. Tesis doctoral. Departamento de Física Aplicada, Facultad de Física, Universidad de Santiago de Compostela. (Septiembre 1994)