

Análisis y paralelización de inteligencia artificial para la reducción de incertidumbre en modelos de predicción de incendios

Pamela Chirino, Germán Bianchini, Paola Caymes Scutari, Mariela Galdamez, Karvin Díaz

LICPaD (Laboratorio de Investigación en Cómputo Paralelo/Distribuido)
Universidad Tecnológica Nacional – Facultad Regional Mendoza

Abstract

En este informe se analizaron métodos de inteligencia artificial, particularmente dos: redes neuronales y visión computacional, con el fin de analizar su posible incorporación al modelo de predicción de incendios llevado a cabo en el laboratorio de Cómputo Paralelo Distribuido(Licpad). Además analizaremos su posible paralelización con el fin de no afectar y si es posible mejorar el rendimiento del sistema actual.

//(Revisar)

Palabras Clave (Times New Roman, 10, negrita)

Inteligencia Artificial, Redes Neuronales, Visión Computacional, Paralelismo

1. Introducción (Times New Roman, 12, negrita)

La inteligencia artificial ha generado una revolución en el mundo de la computación en los últimos años aplicandose a diversos campos. Esta disciplina trata de crear sistemas capaces de aprender y razonar como un ser humano. Entre las ventajas destacables de la misma para nuestra investigación cabe mencionar que son sistemas capaces de analizar grandes cantidades de datos, identificar patrones y tendencias y formular predicciones automáticas con rapidez y precisión.

Con el fin de reducir la incertidumbre en variables y añadir un método de tratamiento de imágenes a modelo de predicción de incendios desarrollado en el Laboratorio de Investigación de Cómputo Paralelo/Distribuido se estudiarán en este informe dos métodos de inteligencia artificial: Redes Neuronales y Visión Computacional; y cómo y con qué propósito se integrarán de manera eficaz al modelo de predicción de incendios.

Comenzaremos mencionando aspectos relevantes del modelo de predicción de incendios y luego detallaremos los dos métodos a tratar y cómo llegarían a impactar en el modelo actual. Además se analizará la forma correcta de paralelizar dichos métodos para obtener una mejora en el tiempo de procesamiento de los mismos

2. Modelo de Predicción de Incendios

Actualmente el modelo de predicción de incendios se trabaja a través del sistema ESS(Evolutionary-statistical System), representado en la figura 1 con las correspondientes entradas y salidas usadas actualmente en el mismo [1].

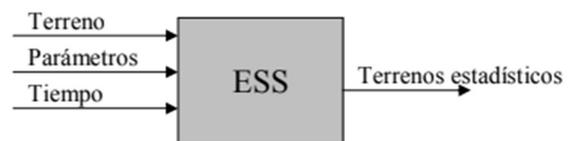


Fig 1: Esquema simple de entradas y salidas del sistema.

El archivo de terreno de entrada se representa a través de una grilla o matriz, donde cada posición indica en qué instante del tiempo fue alcanzado por el fuego dicho sector del terreno a partir de las celdas iniciales. Esta información se representa en un archivo de texto, el cual se ha generado a partir de imágenes tomadas a lo largo de un determinado periodo de tiempo. Este archivo contiene una serie de coordenadas(x,y), que determinan el contorno de cada una de las líneas de fuego.

2.1. Incertidumbre en Variables

Además del archivo de terreno, el sistema de predicción de incendios cuenta con otras

variables de entrada. Ejemplos de estas variables son: dirección y velocidad del viento, el tipo de vegetación, la pendiente de terreno y las dimensiones.

Estas variables suelen presentar cierta incertidumbre debido a que no se pueden medir en tiempo real en el transcurso del incendio.

Actualmente se hace variar estas variables en cierto rango, para obtener una predicción a partir de todas las posibilidades que pueden presentar las variables que desconocemos.

3. Inteligencia Artificial

La Inteligencia Artificial es la simulación de inteligencia humana por parte de las máquinas. Dicho de otro modo, es la disciplina que trata de crear sistemas capaces de aprender y razonar como un ser humano [2].

Por estas cualidades de la inteligencia artificial anteriormente mencionadas es que pretende usarse, para poder obtener una aproximación más exacta de predicciones, tanto analizando la incertidumbre de variables con redes neuronales, como automatizando el análisis y conversión de imágenes a través de la visión computacional.

3.1. Redes Neuronales

Las redes neuronales artificiales son un modelo computacional inspirado en el comportamiento observado en el cerebro humano. Consiste en un conjunto de unidades, llamadas neuronas artificiales, conectadas entre sí para transmitir o comunicar señales. La información de entrada atraviesa la red neuronal (donde se somete a diversas operaciones) produciendo valores de salida [3].

Una red neuronal artificial está constituida por neuronas interconectadas y arregladas en capas, como se muestra en la Figura 2. Los datos ingresan por medio de la “capa de entrada”, pasan a través de la “capa oculta” y salen por la “capa de salida”. Cabe mencionar que la capa oculta puede estar constituida por varias capas de neuronas.

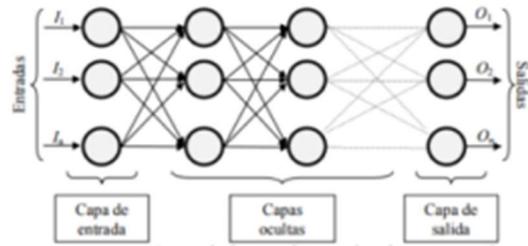


Figura 2: Esquema de una red neuronal.

Aprendizaje en Redes Neuronales

El aprendizaje es el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se producen durante el mismo se reducen a la destrucción, modificación y creación de conexiones entre las neuronas.

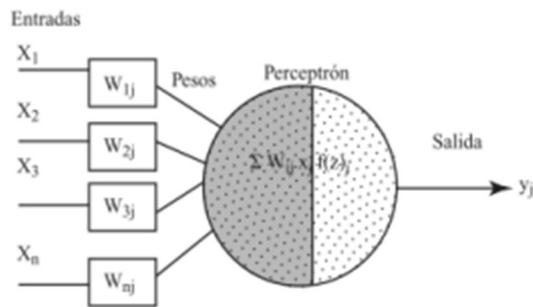
Una red neuronal debe aprender a calcular la salida correcta para cada arreglo o vector de entrada en el conjunto de ejemplos. En los modelos de redes neuronales artificiales, la creación de una nueva conexión implica que el peso de la misma pasa a tener un valor distinto de cero.

El método de aprendizaje que nos interesa para este artículo es el aprendizaje supervisado, que se caracteriza porque el proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo, que podría ser el programador (supervisor, maestro) que determina la respuesta que debería generar la red a partir de una entrada determinada. El supervisor controla la salida de la red y en caso de que ésta no coincida con la deseada, se procederá a modificar los pesos.

Para el estudio práctico de lo presentado anteriormente, estudiamos y ponemos en práctica este tipo de aprendizaje con la red neuronal Perceptrón.

Perceptrón

El perceptrón es la red neuronal más básica que existe de aprendizaje supervisado. El funcionamiento del perceptrón es muy sencillo, simplemente lee los valores de entrada, suma todas las entradas de acuerdo a unos pesos y el resultado lo introduce en una función de activación que genera el resultado final. En la Figura 4 se puede observar un esquema del perceptrón.



Un aspecto importante a tener en cuenta es que el perceptrón utiliza un tipo particular de aprendizaje supervisado llamado Retro propagación de error o backpropagation. Es un método de aprendizaje supervisado con descenso del gradiente. En las redes de retropropagación primero se aplica un patrón de entrada, el cual se propaga por las distintas capas que componen la red hasta producir una salida de la misma. Esta salida se compara con la salida deseada y se calcula el error cometido por cada neurona de salida. Estos errores se transmiten hacia atrás, partiendo de la capa de salida hacia todas las neuronas de las capas intermedias. Cada neurona recibe un error que es proporcional a su contribución sobre el error total de la red. Basándose en este error recibido se ajustan los pesos sinápticos de cada neurona[3].

Para enseñarle a la red neuronal es necesario entrenar un conjunto de datos, el cual consta de señales de entradas asignadas a una denominada “salida deseada”. En el entrenamiento la salida deseada representa la salida esperada para cierto patrón de entrada, es decir, pasamos a la red un conjunto de entradas con la salida que debería producir ese conjunto.

En cada iteración los pesos de los nodos se modifican usando nuevos datos del conjunto para el entrenamiento.

La salida de la red es comparada con la salida deseada. La diferencia entre la salida de la red y la salida deseada se denomina error de la señal. El algoritmo de retropropagación propaga el error de regreso a todas las neuronas, cuya salida fue la entrada de la última neurona. Luego el error se va propagando a las neuronas de capas anteriores, considerando los pesos de las conexiones.

Se considera que la red ha aprendido cuando el error es 0 o un margen próximo al mismo [6][7].

3.2. Visión Computacional

La visión computacional trata de interpretar las imágenes recibidas por dispositivos y reconocer los objetos, ambiente y posición en el espacio [4]. El objetivo al momento de incluir este tema al estudio se basa en la necesidad de trabajar con información proveniente de imágenes de mapas satelitales. A través de la visión computacional proponemos un análisis más rápido y automático de estos mapas con el fin de ayudar a realizar una predicción más precisa.

Etapas en un proceso de visión artificial

En esta sección se detallarán los pasos fundamentales para llevar a cabo la visión computacional.

El primer paso en el proceso es adquirir la imagen digital. Para ello se necesitan sensores y las capacidades para digitalizar la señal producida por el sensor.

Una vez que la imagen digitalizada ha sido obtenida, el siguiente paso consiste en el procesamiento de dicha imagen. El objetivo del procesamiento es mejorar la imagen de forma que el objetivo final tenga mayores posibilidades de éxito.

El siguiente paso es la segmentación. Su objetivo es dividir la imagen en las partes que la constituyen o los objetos que la forman. En general la segmentación autónoma es uno de los problemas más difíciles en el procesamiento de la imagen. La salida del proceso de segmentación es una imagen de datos que, o bien contienen la frontera de la región o los puntos de la misma.

La elección de una representación es solo una parte de la transformación de los datos de entrada. Es necesario especificar un método que extraiga los datos de interés. La parametrización, que recibe también el nombre de selección de rasgos, se dedica a extraer rasgos que producen alguna

información cuantitativa de interés o rasgos que son básicos para diferenciar una clase de objetos de otra.

En el último lugar se encuentra el reconocimiento y la interpretación. El reconocimiento es el proceso que asigna una etiqueta a un objeto basada en la información que proporcionan los descriptores (clasificación). La interpretación lleva a asignar significado al conjunto de objetos reconocidos.

En la Figura 5 se muestra un diagrama para resumir y entender de manera sintética lo explicado anteriormente.

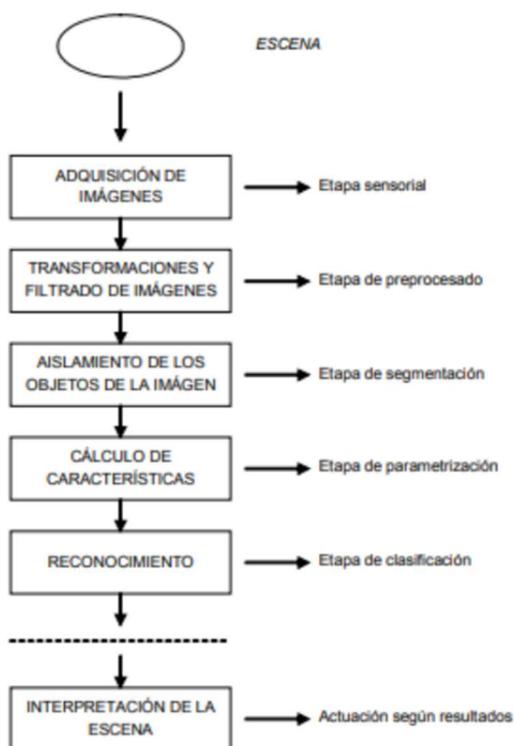


Figura 5: Diagrama

3.2.1. Redes convolucionales

Las redes neuronales convolucionales son muy similares a las redes neuronales ordinarias como el perceptrón multicapa que fue descrito anteriormente; se componen de neuronas que tienen pesos y capacidad de aprender [4].

Lo que diferencia a las redes neuronales convolucionales es que suponen explícitamente que las entradas son imágenes, lo que nos permite codificar ciertas propiedades en la arquitectura;

permitiendo ganar en eficiencia y reducir la cantidad de parámetros en la red.

Las redes neuronales convolucionales trabajan modelando de forma consecutiva pequeñas piezas de información, y luego combinan esta información en las capas más profundas de la red.

Estructura de las Redes Neuronales Convolucionales

En general, las redes neuronales convolucionales suelen estar construidas con una estructura que contendrá tres tipos distintos de capas que se detallan a continuación y se pueden observar en la Figura 3.

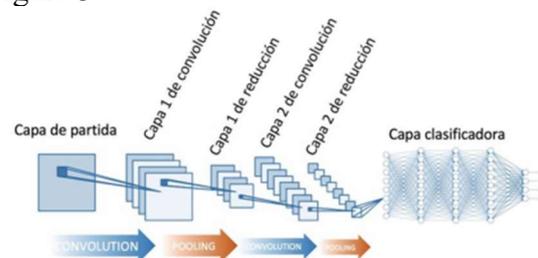


Fig.3: Esquema de la estructura de una red convolucional.

3.2.1.1. Capa convolucional

Esta es la capa que le da nombre a la red, lo que distingue a las redes neuronales convolucionales de cualquier otra red neuronal. En la misma se utiliza una operación

llamada convolución en alguna de sus capas [4].

La operación de convolución recibe como entrada o input la imagen y luego aplica sobre ella un filtro o kernel que nos devuelve un mapa de las características de la imagen original.

De esta forma logramos reducir el tamaño de los parámetros.

3.2.1.2. Capa de reducción o pooling

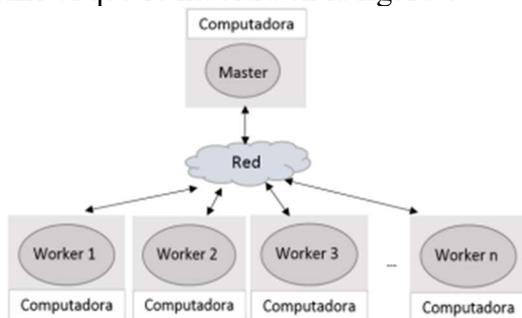
La utilidad principal de la capa de reducción o pooling radica en la reducción de las dimensiones espaciales (ancho x alto) del volumen de entrada para la siguiente capa convolucional [4].

3.2.1.3. Capa clasificadora totalmente conectada

Al final de las capas convolucional y de pooling, las redes utilizan generalmente capas completamente conectadas en la que cada pixel se considera como una neurona separada al igual que en una red neuronal regular [6].

Paralelismo

El procesamiento paralelo es un método mediante el cual se dividen grandes problemas en componentes, tareas o cálculos que puedan resolverse en simultáneo. En particular, estamos analizando el modelo Master-Worker para incluir en nuestro estudio [5]. El modelo Master-Worker es un modelo aplicado a la descomposición de dominio, es decir, el dominio del problema se divide en subconjuntos de datos y los mismos son asignados a nodos diferentes [5], siguiendo un esquema como el que se muestra en la figura 4.



El proceso principal, denominado Master, es el proceso encargado de coordinar todo el tratamiento y procesamiento del problema, para lo que genera muchos subprocesos, que son ejecutados como procesos independientes denominados Workers, que en general se ejecutan en procesadores independientes de forma simultánea. El esquema de trabajo es el siguiente: el proceso Master inicia los procesos Worker, les asigna el trabajo a realizar, y estos devuelven los resultados normalmente en múltiples iteraciones.

// Incertidumbre de variable

En cuanto a aplicar paralelismo en redes neuronales, lo que se propone es paralelizar

el dominio del problema. Cada worker contaría con un perceptrón y estos analizarían las variables con incertidumbre en una x cantidad de casos ejemplo de incendios. Una vez que se dé por finalizado el proceso de aprendizaje de la red neuronal, se analizarían los resultados de exactitud de las predicciones obtenidos con los esperados de los casos de ejemplo.

Como la cantidad de datos a procesar es muy elevada, el Master sería el encargado de enviar el dominio de datos a los diferentes Workers que operarían la red neuronal. De esta manera se entrenaría a los diferentes Workers con los mismos dominios de datos para finalmente analizar el resultado obtenido del aprendizaje, y el correspondiente error que cada uno presente en la predicción. Una vez completado el entrenamiento, se analizarían diferentes combinaciones en este conjunto de variables que permitan obtener el resultado de predicción que esperamos de cada una.

//Visión Computacional

En cuanto a la visión computacional actualmente se está trabajando en la digitalización de las imágenes de los casos de incendios.

Agradecimientos

Si existiera, mencionarlos en forma concisa. Será escrito en fuente (Times New Roman, 10).

Referencias

- [1] EVOLUTIONARY-STATISTICAL SYSTEM (ESS)-Germán Bianchini
- [2] Ponce Cruz P., "Inteligencia Artificial Aplicada a la ingeniería", Alfaomega, 2001.
- [3] Carlos Alberto Ruiz, "Redes Neuronales: Conceptos básicos y aplicaciones", UTN-Facultad Regional de Rosario, 2005.
- [4] Sucar Enrique, Giovanni Gómez, "Visión Computacional", 2008.
- [5] Wilkinson B., Allen M., "Parallel Programming". Pearson. 2005.
- [6] Rivera E., "Entrenamiento de redes neuronales

en algoritmos evolutivos”, 2005.

[7] Matich, D. J., “Introducción a las Redes Neuronales Artificiales”, 2001.

Pamela Agustina Chirino. Universidad Tecnológica Nacional - Facultad Regional de Mendoza.
pamelaachirino@gmail.com