



Área Académica Ingeniería Mecatrónica

Curso:
Inteligencia Artificial

Profesor:
Dr. Ing. Juan Luis Crespo Mariño

Proyecto final
Detección de tamaño de objetos transneptunianos por medio de curvas
de difracción.

Estudiante:
Isaac Porras Navarro

Carné:
201125181

I semestre 2017

Introducción

La astronomía actualmente es un área de la ciencia que se encuentra ampliamente desarrollada y estudiada por los científicos que, día a día tratan de descubrir los secretos que ofrece el cosmos. Observatorio Astronómico Nacional en la Sierra de San Pedro Mártir en Baja California, México; es uno de los lugares donde se desarrolla la astronomía de manera amplia, tanto en su forma teórica como observacional.

Uno de los proyectos que se desarrolla actualmente es el TAOS II (Transneptunian Automated Occultation Survey), el cual consiste en una colaboración internacional para lograr medir la distribución de tamaños de objetos pequeños (aproximadamente 1 km de diámetro) en el cinturón de Kuiper e incluso más lejos, como en la Nube de Oort. Poder obtener estos datos, puede brindar información acerca de la formación y la evolución dinámica del sistema solar, entre otros detalles. Cabe destacar, que este tipo de objetos llamados transneptunianos, no han podido ser detectados por los telescopios actuales, debido a que son bastante oscuros. Por esta razón, el proyecto TAOS II pretende detectarlos mediante ocultaciones, ya que cuando un objeto de este tipo pasa en frente de una estrella, hace que su brillo disminuya, lo que hace que parpadee por cierta fracción de segundo (aproximadamente 0,2 s). [2]

Cabe mencionar, que el proyecto TAOS I obtuvo cero detecciones, debido al mal clima en donde fue ubicado y también se presentaron problemas instrumentales, por esta razón, el proyecto TAOS II fue ubicado en la Sierra de San Pedro Mártir. Este contará con 3 telescopios, los cuales contarán con cámaras especializadas para poder medir el brillo de más de 10,000 estrellas de forma simultánea a 20 Hz. Debido a la cantidad de datos resultante es bastante grande, de aproximadamente 2 TB por noche, de manera que se requieren especialistas para poder centralizar estos datos, además de la utilización de tecnologías de información. [6]

El presente proyecto busca colaborar con TAOS II en la búsqueda de una solución para el análisis y manipulación de los datos. Como se menciona, la cantidad de datos que se producen por noche de observación es tan grande que se requiere la implementación de técnicas computacionales para disminuir el tiempo de análisis y también el error en cuanto a detección de objetos, esto comparado con el TAOS I. Para lograrlo se plantea la utilización de algoritmos de inteligencia artificial, donde se estarán realizando los cálculos de forma continua y permanente en el tiempo sin importar si los telescopios se encuentran en operación o no.

A continuación se mostrarán los procedimientos, análisis y resultados obtenidos con la implementación de una red neuronal para la identificación de tamaños de objetos ocultantes para varias estrellas con tipos espectrales diferentes.

Objetivos

Objetivo general:

Identificar curvas simuladas de patrones de difracción para un estrella ocultada por un objeto de la Nube de Oort por medio de algoritmos de inteligencia artificial.

Objetivos específicos:

1. Diseñar una red neuronal para ser entrenada con los datos simulados enviados por los investigadores del OAN en SPM, México
2. Identificar al menos 1 patrón de difracción con ruido para un tamaño de objeto ocultante para una estrella de un tipo espectral con un error aceptable.
3. Identificar el tamaño de un objeto ocultante para un conjunto de estrellas con diferentes patrones de difracción con ruido a un error aceptable.

Descripción del problema

El proyecto TAOS II es una de ingeniería aplicado al descubrimiento científico astronómico, y como toda obra de ingeniería presenta problemas que resolver. En cuanto a los datos obtenidos por las cámaras de los 3 telescopios, como ya se mencionó, se aproxima que será de 2 TB por noche, y donde se estarán tomando fotografías de más de 10 000 estrellas de forma simultánea, haciendo de esto un reto para los expertos que requieren analizar y procesar esta información.

Una vez que se ha tomado la fotografía (20 fotografías por segundo), se requiere procesarla para obtener las curvas de intensidad de luz recibida de cada estrella, y con esto obtener una señal para toda una noche de observación por cada estrella. Esa señal que se obtiene luego del procesamiento estará afectada por el ruido de los instrumentos que se utilizaron para detectarla. Además, sufrirá de cambios a lo largo de la misma dependiendo de los fenómenos que suceden mientras se realiza la toma de la fotografía. El fenómeno que se desea observar es el de la difracción por el paso de un objeto frente a una estrella.

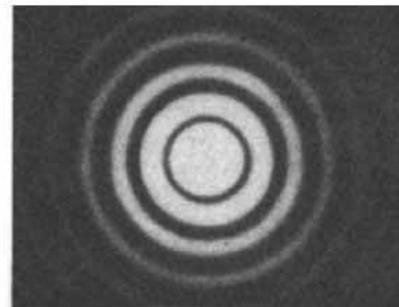
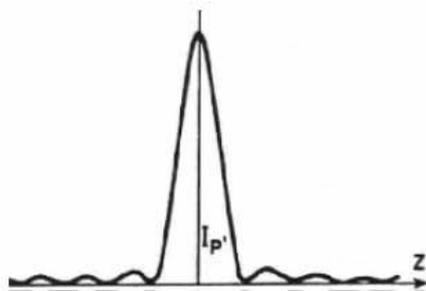
El problema con la información obtenida recae que además de ser gran cantidad a lo largo de una noche de observación, y de estar influenciada por el ruido; va a variar en amplitud dependiendo del tipo espectral de la estrella que se esté observando. Sumado a lo anterior, dependiendo del factor de impacto del objeto sobre la estrella así se producirá el patrón de difracción y por ende así será la señal, dando como resultado una gran variedad de patrones dependientes del factor de impacto. Por último, la curva de difracción que se obtendrá con los telescopios dependerá de la distancia a la que se encuentre el objeto ocultante del observador (la tierra), por lo que todos los patrones que se pueden obtener cuando se da una ocultación es relativamente imposible de analizar por una persona; es aquí donde se vuelve necesaria la implementación de algoritmos computacionales.

Por la variedad de curvas que se pueden obtener se vuelve muy complicado utilizar métodos analíticos para el análisis de la información. También por los niveles de ruido con que se deberá procesar la señal, las funciones de corrección del mismo podrían arrojar valores que dificultan el análisis así como el tiempo de cálculo para todos los procedimientos haría del procesado de la señal una tarea muy compleja. Por esto se plantea el uso de redes neuronales, y que el procesado (identificación de tamaños, cálculo con el factor del ruido, cálculo de tamaños, variaciones en el factor de impacto, tipo espectral y distancia del observador al objeto) de las señal proveniente de cada estrella sea realizada por la red.

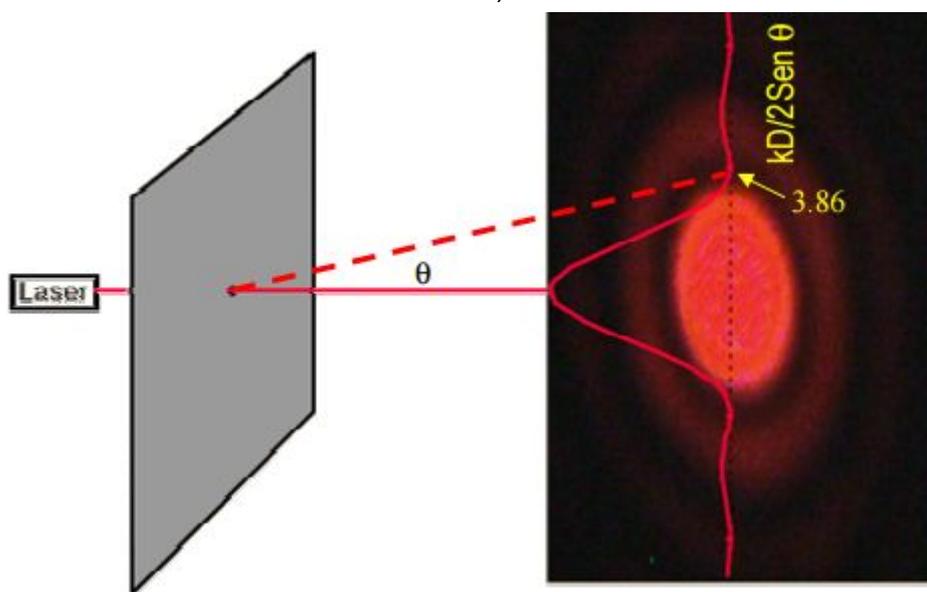
Marco teórico

Patrón de difracción

Al estudiar los patrones de difracción, es de suma importancia estudiar de forma particular el patrón de difracción cuando la luz pasa por una abertura circular, ya que muchos de los sistemas ópticos se componen de lentes y diafragmas circulares. Al pasar la luz por estas aberturas se interrumpe la propagación rectilínea, por lo que aparecen fenómenos más acentuados cuando los obstáculos o las aberturas se hacen más pequeños en relación con la longitud de onda utilizada. [5]. En la Figura 1a se puede observar la forma del patrón de difracción para una abertura circular formada por anillos concéntricos, el círculo central se conoce como Airy y corresponde al orden cero de difracción. Asimismo, en la Figura 1b se puede observar como se determina la representación gráfica de la intensidad en un plano diametral, donde el 90% de la intensidad se ubica en la parte central del patrón de difracción.



a)



b)

Figura 1: a) y b) Difracción por una abertura circular. [5]

Factor impacto

Este es un factor que puede tener un valor de 0 a 1, y se le denomina a la relación que existe entre el paso del objeto por el centro de la estrella y su parte más externa. Este factor corresponde a 1 cuando el objeto pasa por el centro de la estrella, y por su parte, el factor de impacto es igual a 0 si el objeto pasa por la periferia de la misma.

Unidad Astronómica

La Unidad Astronómica (UA) corresponde a una unidad utilizada en el área de la astronomía, es una unidad para expresar las distancias con respecto a las dimensiones del Sistema Solar. Según la Unión Astronómica Internacional, 1 UA se define como "la distancia desde el centro del Sol a una partícula de masa pequeña que, siguiendo una órbita circular, tuviera un periodo de traslación de 365,2568983 días. 1 UA es exactamente 149.597.870,691 kilómetros". [4]

Es decir, que una unidad astronómica es aproximadamente 150 millones de kilómetros, poco menor a la distancia entre la Tierra y el Sol. Cabe mencionar, que estas unidades se comenzaron a utilizar como patrón para poder determinar las distancias de los elementos, ya que no se sabía con precisión la distancia entre la Tierra y el Sol, tan solo un aproximado.

Nube de Oort

La nube de Oort es una zona donde se encuentran los cuerpos pequeños y helados que dan origen a los cometas que llegan a cruzar por la Tierra, esta zona se encuentra más lejana de Plutón. Los cometas tardan miles de años en interactuar con el sistema solar pero cuando pasan cerca de la órbita de los planetas gigantes, estos son atrapados y pueden cambiar su dirección. La predicción de la existencia de esta zona fue de Jan Hendrik Oort, este holandés lo predijo gracias a la observación del paso de los cometas en su época; donde descubrió que los cometas tenían que venir de una zona lejana dentro del sistema solar, en la Figura 2 se muestra el modelo de la nube de Oort que ha sido propuesto actualmente. [3]

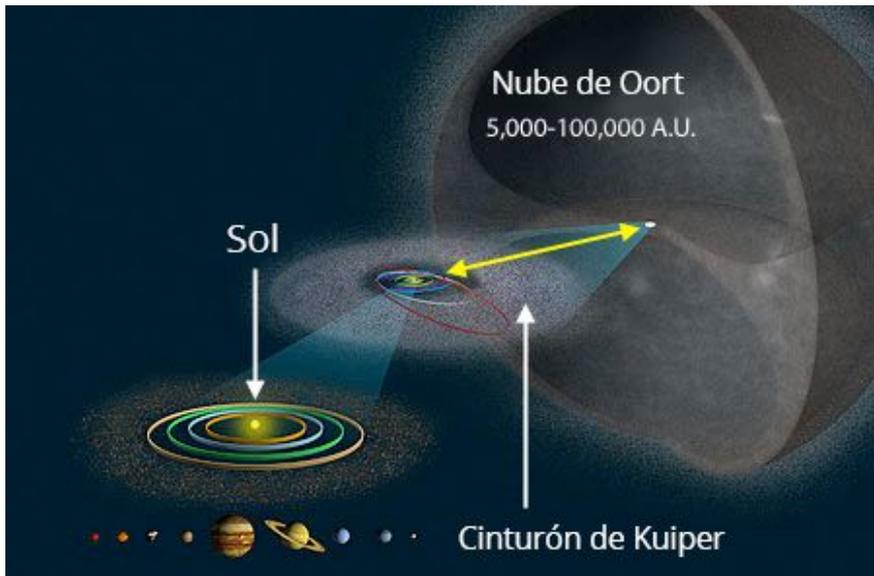
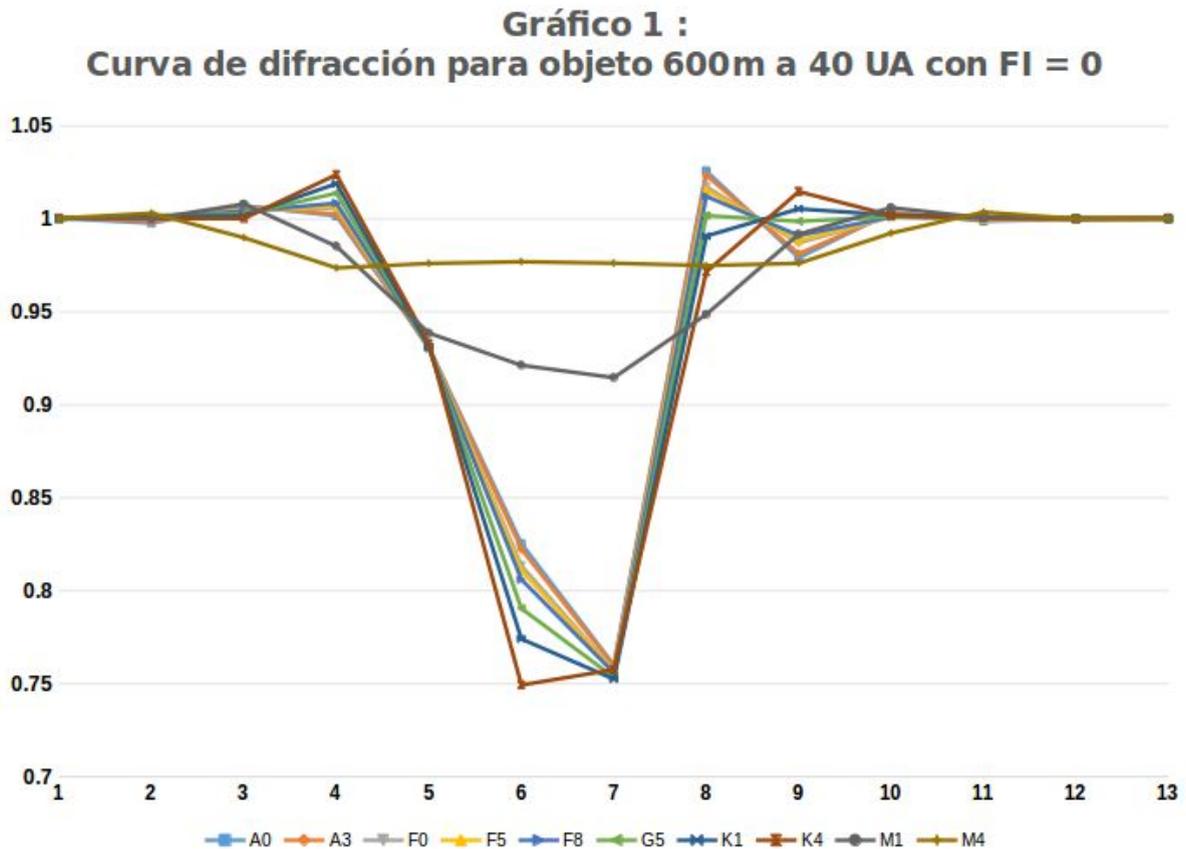


Figura 2: Modelo de la nube de Oort [3]

Descripción de la solución

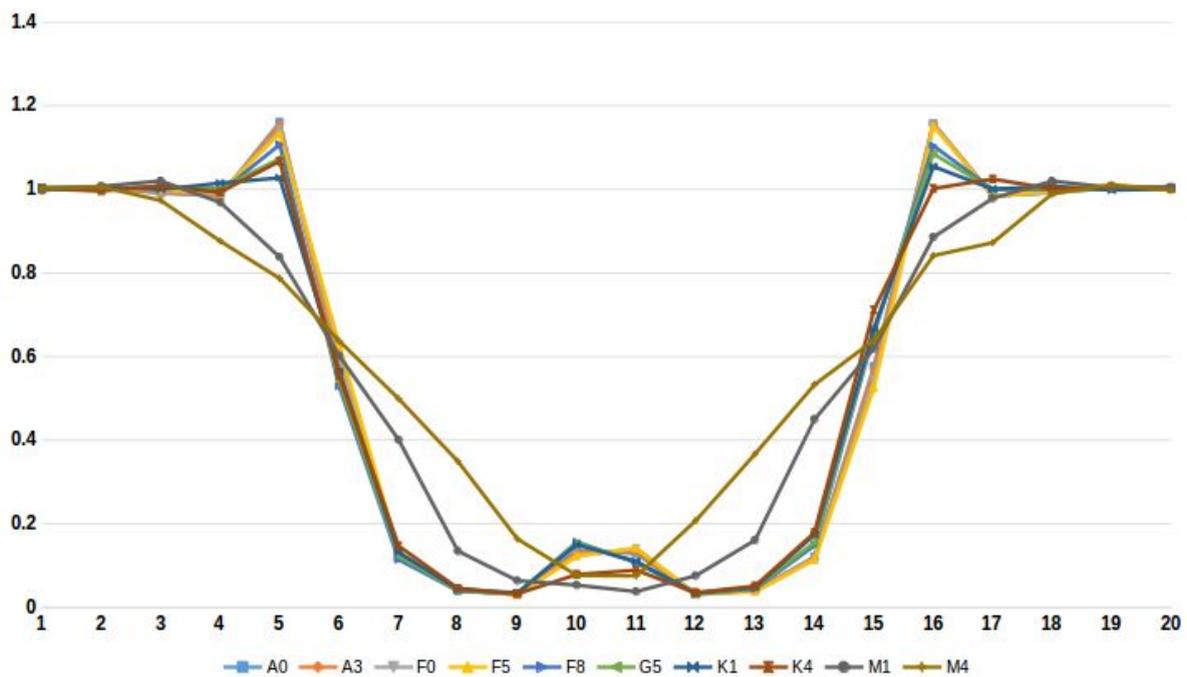
Para el problema descrito anteriormente, en primer instancia y cumpliendo con lo descrito en los objetivos específicos, se diseñará una red neuronal que se adapte a la naturaleza del problema que se posee, y para ello es conveniente hacer un análisis de los datos que se van a utilizar para el entrenamiento. Los valores utilizados originalmente se encuentran normalizados para facilitar la manipulación de los mismos. En los siguientes gráficos solo se mostrará los valores más significativos para el entrenamiento de la red, pues cada simulación consta de 101 datos y para cada curva más de 50 datos son solo la representación de la señal cuando no existe refracción; tampoco se mostrarán todos los gráficos para todos los objetos que se poseen, solo se evaluará el tamaño de objeto menor, el intermedio y el mayor, con base en esto se determinará cómo realizar el diseño de la red.



En el Gráfico 1, se observa la curva de difracción para un objeto ocultante de 600 m a 40 UA de distancia del observador con un FI = 0. Para éste caso, se puede observar que el valor mínimo se encuentra cercano a 0.75, y el ancho del patrón de aproximadamente 8 valores (del 3 al 10); esto sería equivalente a 0.4 segundos de observación en el telescopio. Es importante mencionar que la curva que se produce al ocultar a una estrella para cada tipo espectral que se muestra, es diferente entre sí. Para el tipo espectral K4 se observa una curva más marcada que para el tipo M4.

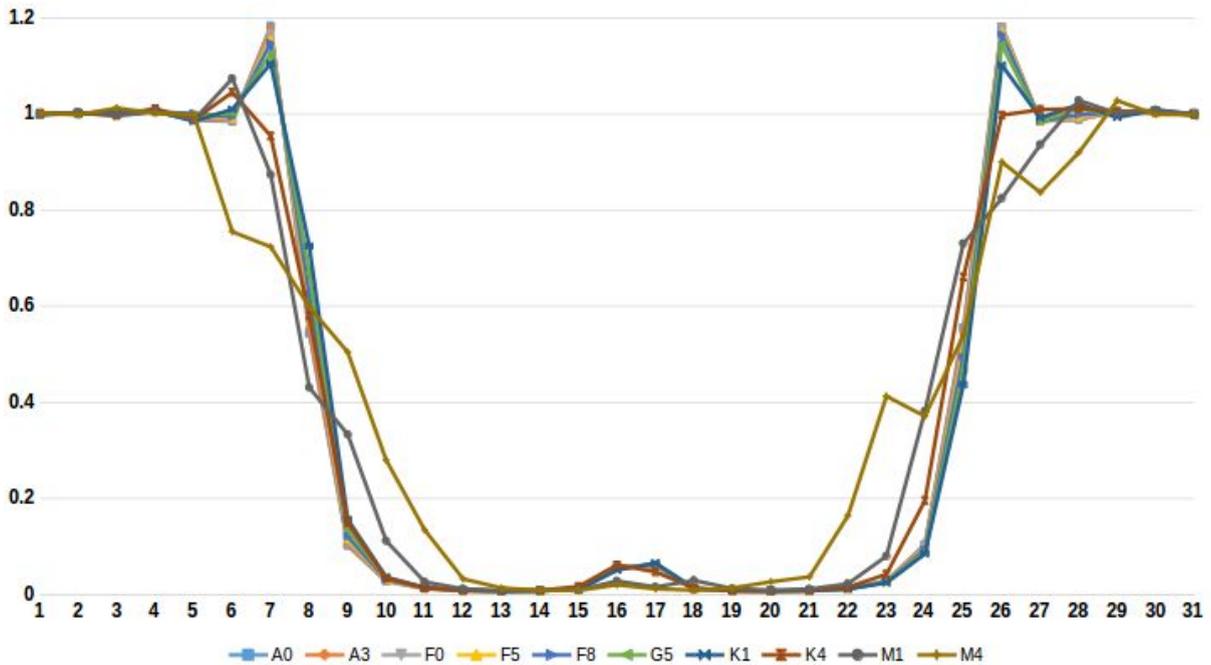
En el Gráfico 2, se observa la curva de difracción para un objeto ocultante de 5100 m a 40 UA de distancia del observador con un FI = 0. Para éste caso, se puede observar que el valor mínimo se encuentra muy cercano a 0, y el ancho del patrón de aproximadamente 16 valores (del 3 al 18); esto sería equivalente a 0.8 segundos de observación en el telescopio. Es importante recalcar que, con respecto al Gráfico 1, tanto la amplitud como el ancho de la curva aumentaron significativamente, por lo que esta característica es importante tenerla en cuenta cuando se diseñe la red que identifique cada uno de los gráficos anteriores.

Gráfico 2:
Curva de difracción para objeto 5100m a 40 UA con FI = 0



En el último gráfico que se analizará para el diseño de la red, Gráfico 3, se observa la curva de difracción para un objeto ocultante de 10100 m a 40 UA de distancia del observador con un FI = 0. Para éste caso, se puede observar que el valor mínimo es 0, y el ancho del patrón de aproximadamente 25 valores (del 5 al 29); esto sería equivalente a 1.25 segundos de observación en el telescopio. Para éste último gráfico, con respecto al Gráfico 2, se acentúa más la amplitud y el ancho del patrón de difracción. También, importante a resaltar es el hecho de que conforme se aumenta el tamaño del objeto ocultante, la curva para cada tipo espectral se van asemejando más entre sí, por ejemplo desde la curva para el tipo espectral A0 hasta el K1 son muy parecidas entre sí, lo que las diferencia son los valores máximos, por lo que esto se vuelve de gran importancia en el diseño de la red neuronal.

Gráfico 3:
Curva de difracción para objeto 10100m a 40 UA con FI = 0



Con lo mencionado anteriormente, se plantea una red neuronal que, tal como plantea la teoría, debe respetar la naturaleza del problema; la red que se propone debe de estar ligada al tiempo, en este caso es al hecho de los datos que forman la curva se encuentran en cierto orden y que el mismo se debe respetar para que el patrón sea formado de forma correcta. Esta dependencia del tiempo, y como se requiere de todos los datos significativos para la identificación de la misma, se utilizará una Red Neuronal de Tiempo de Espera (TDNN por sus siglas en inglés) [1].

Por las características que presentan las curvas antes vistas, se escoge las que se obtuvieron con el objeto de tamaño 10100 m; ya que es la que se determina como más propicia para realizar las primeras pruebas; una vez alcanzados los resultados se implementará la solución en el restante de curvas.

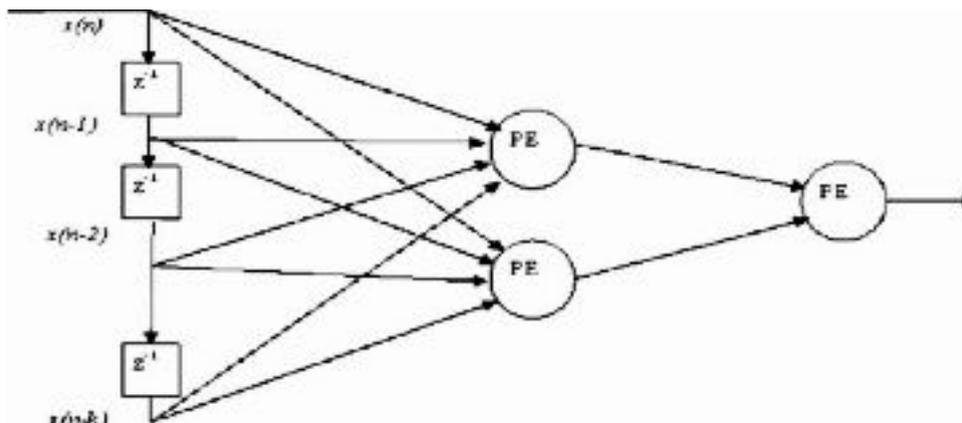


Figura 3: TDNN. [7]

Se propone una red multicapa, con 30 neuronas en la capa de entrada, esto por lo que se menciona con anterioridad de que se requieren los datos representativos de la curva; una capa oculta con al menos 30 neuronas, esto para comenzar con el entrenamiento, si se quieren más se aumentará la cantidad, si con 30 opera de manera conveniente entonces se evaluará si se requieren menos. Por último, se utilizará solo 1 neurona de salida, en la que en primer instancia se utilizará para cumplir con el segundo objetivo, identificar una curva de un tamaño de objetivo determinado y tipo espectral de la estrella específico; y de seguido una vez realizado el aprendizaje para 1 curva esta misma estructura de red se utilizará para identificar un tamaño de objeto para todos los tipos espectrales con que se cuenta.

En cuanto al criterio de parada, se implementarán dos criterios, el primer criterio a evaluar será la cantidad de iteraciones, se sabe que se utilizarán muchos datos, y hasta el momento se desconoce lo que tardará el tiempo de cómputo para los cálculos que se requieren hacer, por lo que éste será el primer criterio. Luego de que se haya obtenido resultados con el criterio anterior, se utilizará el criterio de parada por el error, sin embargo cuánto es el error mínimo al que se entrena la red es una incógnita hasta el momento, dicho valor se propondrá cuando se esté realizando el entrenamiento de la red.

Para la actualización de los pesos se utilizará el método conocido como Regla de Aprendizaje de Corrección de Error (Error-Correction Learning Rule) [1], que se utiliza para el entrenamiento de 1 sola neurona pero extrapolado a una red neuronal. Por otro lado, la función de activación a utilizar será de escalón centrado en 0, y los límites de la misma serán 1 el superior y -1 el inferior. Se utiliza la función escalón pues como lo que se desea es clasificar, por experiencia del diseñador, es la función que mejor se acopla a éste tipo de problema.

Todo lo anterior mencionado se realizará en el software Octave.

Resultados

A continuación se mostrarán los resultados más relevantes que se obtuvieron, se hará de dicho modo ya que para obtener los siguientes resultados se realizaron gran cantidad de entrenamientos, pruebas y variaciones a los parámetros de la red neuronal.

En primer instancia, se debió preparar los datos que iban a ser utilizados para el entrenamiento, por lo que se realizó un proceso largo de organización de los mismos, esto pues se contaban con las simulaciones de los patrones pero sin ningún factor de ruido aplicado, además de que era conveniente en organizarlos por tamaño de objeto y tipo espectral de la estrella. Luego de que se habían organizado, se procedió a incluirlos como variables en Octave. Lo anterior requirió una gran cantidad de tiempo ya que eran muchos los datos que se debían tratar.

Una vez realizado el código para el entrenamiento de la red, se realizó la búsqueda de los valores de los términos de sesgo, α y n (tasa de aprendizaje), con esto lo que se realizó fue modificarlos de forma manual hasta que se visualizará la reducción del error en la red. Una vez encontrados estos valores, se procede a entrenar la red para la identificación de un patrón de difracción para un objeto 10100.

Luego de realizar el entrenamiento en múltiples ocasiones, se logró obtener un resultado aceptable. Antes de ellos, en la mayoría de los casos el error era constante pero no se obtenía ningún dato esperado; en ocasiones se caracterizaba a una curva como la esperada pero al inicio de la misma (cuando los primeros valores significativos de la curva se ingresaban a las neuronas de entrada).

En múltiples ocasiones, cuando actualizaba los pesos, corregía el error para el valor actual y esto hacía que el resultado fuera el deseado, pero inmediatamente lo corregía para el siguiente valor pues en el siguiente valor el resultado ya no era el esperado, observando esto se decidió volver a entrenar la red pero con los pesos obtenidos y un valor n mucho menor a lo que se tenía.

Ahora el resultado había mejorado pero aún no se obtenía un error aceptable, por lo que se realizó un proceso de aumentar la cantidad de neuronas, desde las 30 que se utilizaron inicialmente hasta 500 neuronas, pero el resultado seguía siendo el mismo y lo que se afectaba era el tiempo de ejecución del algoritmo.

Como no se obtenían resultados “buenos”, se procedió a realizar una revisión de los valores que se utilizaban en el entrenamiento, observando que los mismos además de variar por el ruido, para un tamaño de objeto de 9600 m para la estrella de tipo espectral A0 comparado con otra curva de un objeto de 10100 m y del mismo tipo espectral se observó que los patrones eran muy similares; por lo que se decidió ampliar más los datos utilizados para el entrenamiento, incluyendo los valores de 9100 m también.

Luego de múltiples corridas del algoritmo, donde para cuando el error no tendía a disminuir si se le reducía la tasa de aprendizaje y se mantenían los mismos pesos; se logró obtener un red entrenada con los pesos requeridos para que el error no fuera tan alto comparado con los obtenidos hasta el momento. Además, los resultados se obtuvieron con 14 neuronas en la capa de entrada y 10 en la capa oculta.

Gráfico 4: Error vs iteraciones del entrenamiento de la red neuronal

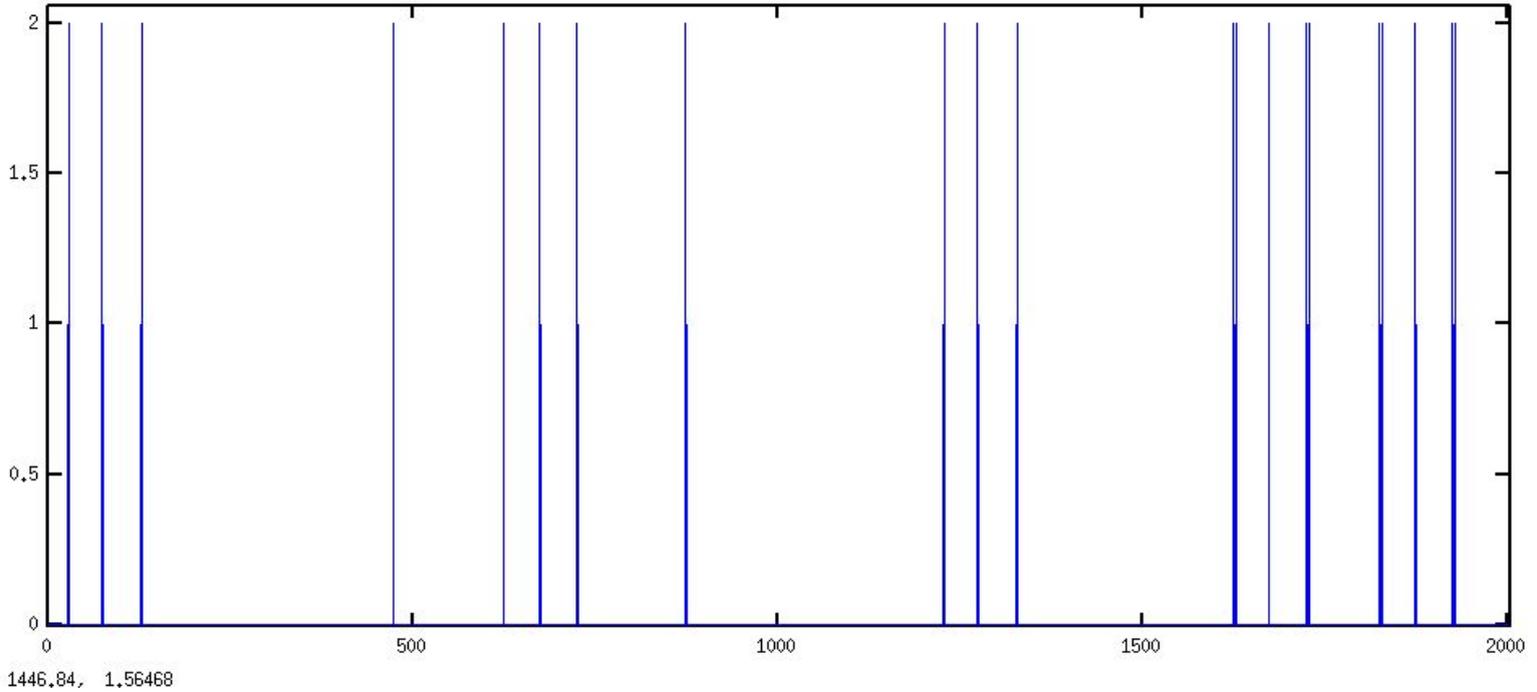
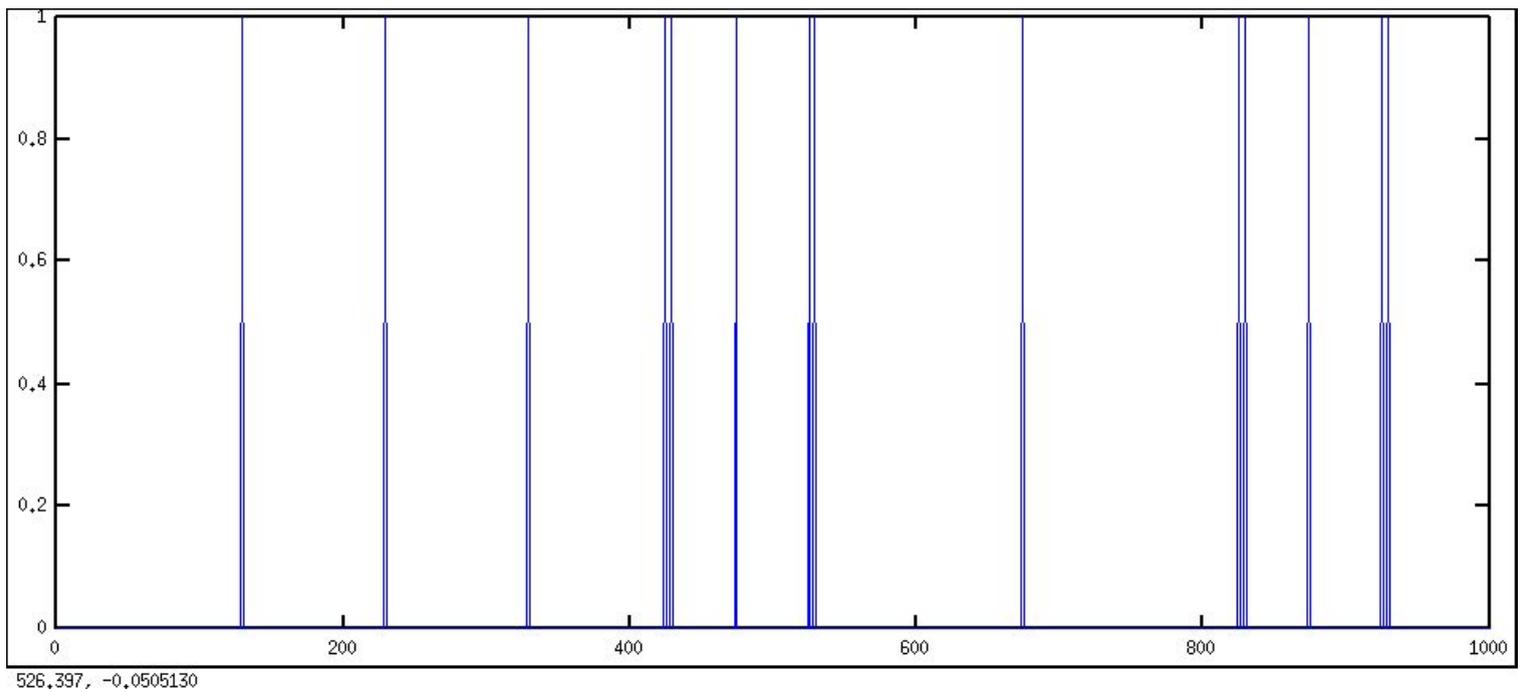


Gráfico 5: Error vs iteraciones del conjunto de prueba de la red neuronal.



Análisis de los resultados

A pesar de la gran cantidad de resultados obtenidos, solo se hará análisis de los cuales el resultado fue o el cambio realizado fue positivo, esto para no hacer extensivo dicho análisis.

En cuanto a las actualizaciones para los pesos, el salto que se producía era evidente que la tasa de aprendizaje era quien lo producía. Esto se debía a que una vez alcanzado un "mínimo local" (se coloca entre comillas pues los mínimos locales son estudiados para el aprendizaje por descenso de gradiente pero su concepto se utilizará de igual modo para éste caso), el salto para seguir disminuyendo en error era muy grande, por lo que los pesos no podían tener ajustes más pequeños y así alcanzar un error menor.

Se menciona en los resultados que, para cierto entrenamiento se obtenía que la red identificaba un patrón de difracción al momento en que los primeros valores de la curva se ingresaban a las neuronas de entrada, esto no se aceptó pues no se puede asegurar que sea un patrón de difracción y no una ocultación por algún otro motivo.

En el Gráfico 1 se observa la cantidad de errores por cada iteración que se daba, es notable que no existen más de 2 errores por cada una, mostrando que el algoritmo se estaba actualizando de la mejor manera y su error estaba minimizando, a pesar de no disminuir conforme pasaban las iteraciones, el resultado fue aceptable. Existe una cierta periodicidad en el error, esto se debe a la organización de los datos con que se estaba entrenando la red.

En el Gráfico 2, se hace notable que se obtuvieron más errores cuando se realizó la prueba a los pesos y así verificar que fueran aceptables. A pesar de que el error se mantenía, de igual forma que en el aprendizaje el resultado que se obtuvo fue mejor que en los anteriores. En algunos casos existe un error porque la red clasificó la curva como la que se esperaba pero en un punto de la señal donde no se le había indicado que lo debía de hacer, pero al encontrarse tan cercanos los valores el error se acepta.

No se realizó un entrenamiento para el segundo objetivo por la dificultad que presentaba y porque no se contaba con el tiempo que se requería para hacerlo.

Conclusiones

- Se diseñó una red neuronal que puede ser entrenada con los datos simulados enviados por los investigadores del OAN en SPM, México.
- Se identificó 1 patrón de difracción con ruido para un tamaño de objeto ocultante para una estrella de un tipo espectral con un error aceptable.
- No se identificó el tamaño de un objeto ocultante para un conjunto de estrellas con diferentes patrones de difracción con ruido a un error aceptable.
- Taos II es proyecto ingenieril muy complejo, que requiere de la implementación de herramientas computacionales avanzadas para asegurar su éxito.
- Las redes neuronales son una herramienta muy útil para el procesado de gran cantidad de datos que se encuentran afectados por múltiples variables.

Recomendaciones

- ★ El aprendizaje hacerlo con un sistema de cómputo de alto desempeño, para reducir los tiempos que el algoritmo tarda en aprender y disminuir su error.
- ★ Evaluar funciones de activación distintas al escalón, como lo son la sigmoide y la rampa, esto para la implementación del cálculo del tamaño del objeto.

Referencias

- [1] Engelbrecht; A. (2007). *Computational Intelligence An Introduction*, 2nd ed. England.
- [2] Reyes; M. & Castro, J. (2016). *Retos Científicos y Tecnológicos del Proyecto TAOS II*.
- [3] Robles, T.. (2015). CONACYT, Modelo de la nube de Oort. URL: <http://www.conacytprensa.mx/index.php/ciencia/universo/3212-proponen-modelo-mexicano-de-la-nube-de-oort> , 2015. [En línea; Consultado en 10 Jun. 2017].
- [4] SEA. Sociedad española de astronomía, Unidad astronómica. URL: <http://www.sea-astronomia.es/drupal/node/342>. [En línea; Consultado en 10 Jun. 2017].
- [5] UCM. Universidad Complutense Madrid. Tema 6. Difracción. URL: http://pendientedemigracion.ucm.es/info/opticaf/OPT_FIS/apuntes_sueltos_prov/pdf/Tema6.pdf. [En línea; Consultado en 10 Jun. 2017].
- [6] UNAM. (2016). Proyecto TAOS II. URL: <http://www.astroscu.unam.mx/taos/index.php/en/>, 2016. [En línea; Consultado en 10 Jun. 2017].
- [7] Yonas, B. (2006). Temporal neural networks for downscaling climate variability and extremes. URL: https://www.researchgate.net/figure/7251784_fig1_Fig-1-Schematization-of-the-Time-lagged-feed-forward-neural-network-TLFN-with-one , 2016. [En línea; Consultado en 10 Jun. 2017].

Anexos

Anexo 1: Algoritmo para entrenamiento de la red neuronal.

El algoritmo se encuentra en un archivo adjunto, se identifica con el nombre de Red_esc_10100.m

Anexo 2: Algoritmo para prueba de la red neuronal.

El algoritmo se encuentra en un archivo adjunto, se identifica con el nombre de Red_esc_10100_prueba.m

Anexo 3: Pesos obtenidos del entrenamiento de la red neuronal.

Los pesos se encuentran en un archivo adjunto, los de la capa oculta se identifican con el nombre de v_bueno y el de la capa de salida con w_bueno.