

**Desarrollo de un sistema de  
identificación de sistemas de riego de  
pivote central basada en una red  
convolucional 3D e imágenes satelitales  
de alta resolución.**

**Tutor: Kernel Enrique Prieto Moreno**

**Alumno: Arturo Jair Mazón Ocampo**

Cuarto Verano de la Investigación en Matemáticas 2019

Unidad Cuernavaca del Instituto de Matemáticas,  
UNAM  
Universidad Autónoma del estado de Guerrero

# Contents

<b>1</b>	<b>Introducción</b>	<b>1</b>
1.1	Antecedentes . . . . .	1
<b>2</b>	<b>Objetivos</b>	<b>2</b>
2.1	General . . . . .	2
2.2	Especificos . . . . .	2
<b>3</b>	<b>Hipotesis</b>	<b>2</b>
<b>4</b>	<b>Metodologia</b>	<b>3</b>
4.1	Redes convolucionales 3D . . . . .	3
4.1.1	Redes neuronales artificiales . . . . .	3
4.1.2	Componentes de una red neuronal . . . . .	3
4.1.3	Red Neuronal Convolutiva . . . . .	3
4.1.4	Red Neuronal Convolutiva 3D . . . . .	4
4.2	Elaboracion del banco de imagenes . . . . .	4
4.3	Sistema de indentificacion de riego de pivote central basado en una 3DCNN . . . . .	5
4.3.1	Preprocesamiento de datos . . . . .	5
4.3.2	Informacion de entrenamiento . . . . .	5
4.3.3	Arquitecturas de la 3DCNN . . . . .	6
4.3.4	Arquitectura basada en LeNet . . . . .	6
4.3.5	Arquitectura basada en AlexNet . . . . .	7
4.3.6	Arquitectura basada en VGGNet . . . . .	7
4.4	Transformada de Hough circular . . . . .	7

## 1 Introducción

La necesidad de economizar recursos de diverso tipo como personal de trabajo, inversión económica, o simplemente disminuir el tiempo en que ciertas tareas se realizan ha llevado al desarrollo de diversas tecnologías en el campo del machine learning, en particular la rama del deep learning , todo con el fin de llegar a la automatización de varias actividades con el objetivo del ahorro de los recursos antes mencionados. Con lo que respecta al campo Mexicano, el aproximamiento para lograr el ahorro y la eficiencia que necesita el sector agrícola en México es el de optar por el uso de las técnicas de deep learning y el uso de sensores remotos, estos representados por el uso de imágenes satélites de alta resolución en diferentes bandas espectrales,y por la parte del deep learning el uso de redes neuronales convoluciones,por sus siglas en ingles, (CNN) para tratar diversos aspectos, tales como la identificación del tipo de cultivos, el calculo de la evapotranspiracion en una zona dada, o la identificación del área delimitada por los campos de cultivo, en este trabajo nos enfocaremos en utilizar estas herramientas principales que son las CNN y las imágenes satélites de alta resolución para la identificación de campos de cultivo de pivote central.

### 1.1 Antecedentes

El uso de la tecnología de deep learning en conjunción con el uso de imágenes satélites de alta resolución es una tecnología relativamente nueva, teniendo su auge de trabajo aproximadamente en el año 2016 [1],pero que ha intentado profundizar el conocimiento que tenemos sobre las técnicas y redes diseñadas hasta la fecha, ya que uno de los problemas con los que mas se encuentra el deep learning con el estudio de remotes sensores, tales como las imágenes satélites de alta resolución, es la magnitud tanto de cada muestra como de el volumen de las muestras

disponibles , en ese sentido las redes neuronales mas conocidas tienden a tener un menor desempeño en una imagen de alta resolución que de menor resolución, y la otra parte haciendo referencia al hecho de que aun siendo lanzados en 2014 al menos a la fecha de 2017 los satélites Sentinel había recolectado aproximadamente 25 petabytes de información en imágenes satélites de alta resolución[1].

La clasificación de imágenes hiperespectrales, en [2] demostró que varias condiciones atmosféricas. la distribución de la luz , la similitud de intercalases y la variabilidad dentro de las mismas clases resulto en que el análisis hiperespectral resultaba ser no lineal. Esto implica que la adición de la información de bandas hiperespectrales, usualmente infrarrojas, permite añadir mas información a la red neuronal para poder tener una conjetura sobre parametros diferentes a los generados por los canales R,G,B.

El uso de CNN supervisadas: recientes avances en el desarrollo de 3D CNN , las cuales son redes neuronales con un eje tridimensional el cual hace referencia al tiempo, tal arquitectura se ha utilizado en clasificación hiperespectral , en otras palabras, en las 3D CNN las operaciones convolucionales se realizan de manera espacio-temporal, en contraste con las 2D CNN que lo realizan solo de manera espacial. Los autores de [3] introdujeron un modelo basado en una 3DCNN , superbisada, con una regularización l2 y en [4] se uso el modelo de manera exitosa para la clasificación espacio espectral

## 2 Objetivos

### 2.1 General

El objetivo general de este trabajo es crear un modelo basado en Deep learning para la identificación de sistemas de riego de pivote central con la ayuda de imágenes satélites de alta resolución multispectrales que mejore el estado del arte actual.

### 2.2 Especificos

- Elaborar un banco de imágenes para una zona agrícola que contenga sistemas de riego de pivote central de la zona de cultivo de los estados del norte del país i.e. Sonora, Chihuahua etc. y de la zona noroeste de Colorado usando el satélite Landsat.
- Crear un sistema basado en la tecnología imfindcircles de Matlab, y de variación espectral para localizar los SRPC y sus respectivos centros.
- Entrenar tres redes neuronales convoluciones 3D para su clasificación y encontrar la que produzca el mayor desempeño
- Elaborar una clasificación de los sistemas de riego de pivote central entre otros de las imágenes tomadas del norte del país para su eventual comparación con la red convencional 3D entrenada con el crop data layer del noroeste de Colorado.

## 3 Hipotesis

Mediante el uso de una red neuronal convolucional 3D (3DCNN) junto con el uso de imágenes satélites de alta resolución divididas en 4 bandas espectrales: Rojo,Verde, Azul e Infrarojo,por sus siglas en ingles, (RGBI), es posible aumentar la eficiencia de vanguardia para la clasificación de sistemas de riego de pivote central(SRPC), comparado con el trabajo de Chenxiao Zhang et al [5]

## 4 Metodología

### 4.1 Redes convolucionales 3D

#### 4.1.1 Redes neuronales artificiales

Una red neuronal artificial, por sus siglas en inglés, (ANN) es una red conformada por elementos llamadas neuronas artificiales, las cuales reciben ciertos valores de entrada a los cuales se les aplica un proceso, el cual dado cierto criterio permite crear un valor de salida dependiente del valor de entrada y del criterio dado.

La red forma conexiones al conectar los valores de salida de ciertas neuronas a los valores de entrada de otras neuronas, a estas conexiones se les asigna valores a las que llamamos pesos. Los pesos y el criterio con el cual usamos para designar si un valor de salida se produce o no, a este criterio se le asigna el nombre de activación, pueden ser modificados a través de un proceso llamado aprendizaje, el cual está gobernado por una regla de aprendizaje.

#### 4.1.2 Componentes de una red neuronal

- Neuronas

Una neurona con etiqueta  $j$  que recibe el valor de entrada  $p_j(t)$  de neuronas anteriores consiste de los siguientes elementos:

- Una activación  $a_j(t)$ , que depende de un parámetro discreto.
- Un umbral  $\theta_j$ , el cual se queda fijo a menos que sea alterado por una función de aprendizaje
- Una función de activación  $f$  que calcule la activación para el valor  $t + 1$  dado  $a_j(t), p_j(t), \theta_j$ , dando pie a la siguiente relación:

$$a_j(t + 1) = f(a_j(t), p_j(t), \theta_j)$$

- Una función de salida  $f_{out}(a_j(t))$  que calcule el valor de la activación.

$$o_j = f_{out}(a_j(t))$$

- Conexiones, pesos y sesgos

La red consiste en conexiones, en donde el valor de salida de la neurona  $i$  como valor de entrada  $j$ , en ese sentido decimos que  $i$  es predecesor de  $j$ . A cada conexión se le asigna un peso  $w_{ij}$ . A veces términos de sesgo se añaden a los valores de salida para añadir un sesgo hacia el criterio de activación.

#### 4.1.3 Red Neuronal Convolucional

Los avances y trabajos más tempranos relacionados con las ANN llevaron a crear procesos que superaran los estándares de desempeño en ciertas tareas, en específico la clasificación, localización de imágenes, estas necesidades llevaron a la creación de las redes neuronales convolucionales, por sus siglas en inglés, (CNN).

La característica principal de las CNN es que mientras otros procesos intentan encontrar características globales, por ejemplo intentar encontrar un patrón que surja a partir de todos los píxeles de una imagen, las CNN buscan características locales en la imagen.

Esta característica principal trae consigo dos propiedades muy importantes

- *Los patrones que aprende son invariantes a la traslación.* Cualquier patrón que aprenda la CNN siempre podrá detectarla sin importar si este se encuentra en las esquinas, en el

centro, o cualquier otra región de la imagen. Mientras que otro tipo de redes tendrían que aprender el patrón nuevamente si este cambiara de posición, esto permite aumentar la eficiencia cuando se trata de imágenes.

- *La CNN puede aprender jerárquicas con respecto a las características espaciales.* Un ejemplo de esto sería una capa convolucional que aprenda pequeñas características locales como bordes, mientras que la siguiente capa convolucional tomaría estas características como valores de entrada y aprendería características locales más complejas como curvas cerradas, y así sucesivamente.

Las convoluciones operan sobre tensores 3D, llamados mapas de características, con dos ejes espaciales (alto y ancho) así como un eje de profundidad (también llamado eje de canales). La operación convolucional extrae parches a su mapa de entrada y aplica la misma transformación a cada uno de estos parches, este proceso crea un valor de salida que es un mapa de características. Este valor de salida sigue siendo un tensor 3D, tiene altura y ancho, y una profundidad arbitraria, los diferentes canales en el eje de la profundidad representan filtros. Estos filtros extraen características específicas presentes en la información de entrada.

#### 4.1.4 Red Neuronal Convolucional 3D

Una red neuronal convolucional 3D, por sus siglas en inglés, (3DCNN), es una red neuronal donde se agrega un eje extra para la información de entrada, en este caso este eje representa el tiempo, en donde las imágenes son tomadas en la misma región en diferentes tiempos, esto con el propósito de encontrar características espacio temporales [10], las cuales son de gran relevancia para el análisis de vegetación, y en este caso con el análisis de cultivos, la estructura del tensor de los valores de entrada es un tensor 4D, donde la dimensión extra corresponde a la misma imagen en un periodo diferente de tiempo.

## 4.2 Elaboración del banco de imágenes

Considerando el hecho de hacer un estudio comparativo recopilaremos una base de datos con características parecidas a la utilizada por Chenxiao Zhang et al [5], tomando en cuenta también un clima similar al de los estados del norte, tales como: Sonora, Chihuahua, etc. Tomando en cuenta esto se recopilara una base de imágenes a partir del uso de imágenes Landsat y Crop Data Layer (CDL), utilizando la información del satélite Landsat 5-TM descargada de la página del sistema de vigilancia geológica de E.U.A (U.S. Geological Survey) [7], elegimos las imágenes de la zona del noroeste de Colorado por su gran cantidad de sistemas de riego de pivote central (CRPC) con una cobertura de de  $20,000 \text{ km}^2$ , usando las bandas RGB, que coinciden con la Banda 1 (Azul), Banda 2 (Verde), Banda 3 (Rojo), además de una cuarta banda que será parte clave para nuestro estudio que será la banda 5 (Infrarojo), serán la información de entrada de nuestra red neuronal para su entrenamiento.

Con respecto a la información de la distribución de los SRPC en cada imagen, es decir, el CDL fue producido por El servicio nacional de estadística agrícola (NASS) del departamento estadounidense de agricultura (USDA), ofrece una clasificación de varias categorías de cultivos de cobertura en resolución de 30m, con un rango exactitud de entre 85% a 95%. Dado el alto grado de exactitud de la clasificación del uso de suelo, CDL es incorporado en el trabajo para diferenciar áreas sin cultivar en las imágenes Landsat. La información de CDL se puede conseguir de manera gratuita a través de los servicios de la página web CropScape.

Considerando los tiempos de cultivo se eligen 3 conjuntos de imágenes para utilizar la característica temporal de la 3DCNN con un intervalo de aproximadamente una década entre cada conjunto de imágenes, cada conjunto conformado por 3 imágenes tomadas en el mismo año en

el periodo de Abril, la época en donde comienza la plantación de los cultivos, Julio donde los cultivos llegan a la madurez de su crecimiento, y el finalmente Noviembre donde termina el ciclo de crecimiento con el fin de la cosecha de los cultivos.

### 4.3 Sistema de indentificacion de riego de pivote central basado en una 3DCNN

Siguiendo una idea análoga a [5] el trabajo se dividirá en 2 principales etapas, crear una 3DCNN para la clasificación de los SRPC, y abordar el problema de la localización del centro de los SCRP. En particular, para el segundo punto, haremos uso además del sistema de varianza en la información dada por los canales RGB, en donde además de utilizaremos esta misma idea para el canal infrarrojo, encontramos en el trabajo de Hou, W. et al [6] la justificación para incluir el uso de una varianza basada en la banda infrarroja, tenemos que la vegetacion húmeda tienen un rango definido y distintivo de reflexion para las bandas infrarrojas, que ayuda a diferenciar de terreno con agua , suelo desnudo, o vegetación seca.

Esto quiere decir que esta información adicional para la identificación nos ayudara en 2 sentidos, el primero aumentar la habilidad de la red para diferenciar secciones que tengan cultivo de las que no, en el segundo tenemos que dada su estructura, la distribución de humedad de un SCRP es diferente a la de otras distribuciones de cultivo, dado que la humedad es bastante uniforme en toda la zona que abarca el SRPC.

Para crear a los conjuntos de entrenamiento, validación y prueba, se planea cortar representantes de SCRP de cada conjunto de imágenes de la misma manera, cortando el mismo SCRP de cada imagen del conjunto para formar un tensor de la forma (t,w,h,c) donde cada eje representa el tiempo, el ancho , alto y los canales espectrales respectivamente.

Para facilitar la selección manual de los SRPC y sus respectivos centros utilizaremos la información que aporta el CDL para clasificar del terreno con cultivo del que no tiene, para crear conjuntos de SRPC para el entrenamiento de la 3DCNN.

Despues usaremos un sistema de grises con respecto a centros identificados manualmente, en donde el centro del SRPC sea el punto con menor varianza con respecto a todos los demás puntos, además de aunado con un sistema similar con el uso de la banda 4 de infrarrojos con la consideración de usar un filtro extra en donde la varianza tiene que estar en un intervalo especifico , en donde esto facilitara el proceso de clasificación con respecto a otros sistemas de cultivo que podría clasificar incorrectamente una CNN de 3 canales.

#### 4.3.1 Preprocesamiento de datos

Tres canales visibles (banda 1-Azul, banda 2-Verde, banda 3-Rojo) y un canal infrarrojo (banda 5- Infrarrojo) del Landsat 5-TM fueron usados como la principal informacion de entrada para la clasificación.Despues se hacen 2 imágenes una de color utilizando los 3 canales visibles y una usando solo el canal infrarrojo, para localizar manualmente los centros ,siendo la imagen de calor la mas útil para las épocas de finales y principios del proceso de cultivo y la infrarroja para la etapa de maduración.Un área que abarque 300 SRPC sera utilizada como área de entrenamiento, de la cual se obtendrán los conjuntos de entrenamiento y validación, recortada de las imágenes Landsat.

Una imagen de entramiento de área de  $21,000 \text{ km}^2$  la cual usando CDL fue filtrada de terrenos tales como depósitos de agua al aire libre, zona urbana o industrial, y cualquier otro terreno que no fuera de cultivo, lo cual disminuye la gran cantidad de información de entrada.

#### 4.3.2 Informacion de entrenamiento

Para preparar la información que se usara para entrenar nuestra 3DCNN se usaran cortes sobre las imágenes de  $34 \times 34$  píxeles, de donde se tomaran muestras que muestren tanto SRPC como

terreno de otro tipo sea de cultivo o no.

La razón para utilizar cortes de 34x34 esta dada por el hecho de que la mayoría de los SRPC tienen una cobertura circular similar de 27x27 píxeles, esta elección afecta directamente en el desempeño para la clasificación, ya que recortes demasiado grande provocarían la posibilidad de tomar 2 SRPC en el mismo recorte, llevando a una clasificación errónea, mientras que si se toma un corte demasiado pequeño provocaría que características esenciales de los SRPC podrían no ser capturadas, siendo el principal su área circular, e.g. secciones de terrenos de cultivo poligonales demasiado grandes podrían ser clasificados como SRPC. Para cada píxel identificado como el centro de SRPC se uso una ventana de 34x34 que cubra todo el SRPC tomado de la imagen de color para el entrenamiento de la clasificación, y para cada píxel que no se identifica como el centro de un SRPC se toma una ventana de 34x34 alrededor para el entrenamiento para discernir entre un terreno conteniendo un SRPC del que no.

Toda la información de entrada fue etiquetada como SRPC o no-SRPC, pero dada la distribución de estas formas de cultivos y el área de la que se tomo la información de entrada la cantidad de información etiquetada como SRPC era menor que el resto, para lidiar con ese problema se proponen técnicas como transformaciones de imágenes, tales como rotaciones, reflexiones y "cropping" [8, 9].

El método elegido para este trabajo es el de crear muestras a partir de una muestra etiquetada como SRPC en donde se caracterizan por ser la muestra original con el centro movida ligeramente de su posición en alguna dirección.

Para cada SRPC una ventana de 34x34 píxeles centrada con respecto al centro del mismo fue usada para moverse ligeramente de manera horizontal y vertical sobre la imagen original para generar nuevas muestras con el centro del SRPC desviado ligeramente en alguna dirección con respecto al centro de la ventana de 34x34 píxeles, utilizando esta estrategia haciendo 25 movimientos, 5 píxeles horizontalmente y 5 píxeles verticalmente podemos crear 25 veces la cantidad de muestras disponibles para el entrenamiento.

### 4.3.3 Arquitecturas de la 3DCNN

Se entrenaran tres 3DCNN con arquitecturas basadas en LeNet, AlexNet y VGGNet, para comparar sus respectivos desempeños y además comparar los desempeños con respecto a las redes CNN basadas en las mismas arquitecturas que tienen el mismo fin que las redes construidas en este trabajo.

Los valores de entrada para las 3 redes son conjuntos de 3 imágenes de la misma área en 3 periodos de tiempo distinto con 4 canales RGBI, donde cada canal entra de manera independiente en la 3DCNN, cada imagen con un tamaño de 34x34x4 píxeles. Los valores RGBI fueron normalizados para cada banda. Los valores de salida son una clasificación binaria del terreno que presentaba los sistemas de riego de pivote central del terreno que no.

### 4.3.4 Arquitectura basada en LeNet

Esta red esta basada en la arquitectura de LeNet. La red consiste en dos capas convolucionales, dos capas de pooling promedio y dos capas totalmente conectadas. La primera red convolucional utilizo veinte filtros de tamaño 3x3x3, seguido de una operación una operación de rectificación lineal, por sus siglas en ingles, (ReLU) un pooling promedio, que toma el valor promedio de regiones de 1x2x2. Los valores de salida de la primera capa de pooling promedio es un tensor de 3x17x17x20, luego este se proceso en la segunda capa convolucional con cincuenta filtros de tamaño 3x3x3. Se utilizo ReLU despues y seguido de esto se utilizo la segunda capa de pooling promedio con los mismos hiper parametros que la anterior. Los mapas de características fueron convertidos en un solo vector con una función de aplanado. Finalmente este vector fue procesado en una capa totalmente conectada de quinientas neuronas, y a continuación otra capa totalmente conectada de dos neuronas, seguido de una función ReLU y un operador "soft-max".

### 4.3.5 Arquitectura basada en AlexNet

La arquitectura basada en AlexNet es similar a la de LeNet, pero consta de una mayor cantidad de capas, esta arquitectura consiste en cinco capas convolucionales cada una seguida de un operador ReLU, tres capas de pooling promedio, y tres capas totalmente conectadas. Veinte filtros de tamaño 3x10x10 se usaron en la primera capa convolucional, el número de filtros se duplicó a 40 en la segunda capa convolucional de tamaño 20x5x5. Cada capa convolucional fue seguida de una capa de pooling promedio con un tamaño de pool de 1x2x2. La imagen fue convertida en un mapa de características de tamaño 3x8x8x40, después de el proceso de las capas de pooling promedio. El valor de salida de 3x8x8x40 fue después procesado por tres capas convolucionales. La tercera capa contiene ochenta filtros de tamaño 8x8x40, seguido de una capa ReLU. La cuarta capa contiene ochenta filtros de tamaño 8x8x80, seguido también de una capa ReLU. Una configuración de cuarenta filtros de tamaño 8x8x80 en la última capa convolucional. Un pooling promedio se utilizó a continuación de tamaño 1x2x2 en el mapa de características. Finalmente las características de tamaño 3x4x4x40 fueron aplanadas a un vector para usar como valores de entrada para las tres capas totalmente conectadas, las 2 primeras capas totalmente conectadas contienen 500 neuronas, y la última capa totalmente conectada contiene 2 neuronas, que representa la clasificación de si la imagen es un SRPC o no.

### 4.3.6 Arquitectura basada en VGGNet

La arquitectura está constituida en trece capas convolucionales, cinco capas de pooling promedio y tres capas totalmente conectadas. Toda la arquitectura puede pensarse como un proceso de cinco etapas, cada etapa con una configuración similar. Las primeras dos etapas tienen 2 capas convolucionales y una capa de pooling promedio. Tres capas convolucionales y una capa de pooling promedio es la estructura de las siguientes tres etapas. El número de filtros en cada etapa eran duplicados en la siguiente etapa, de 32 hasta 256. Las capas convolucionales usadas en la quinta y cuarta etapa fueron las mismas, filtros de tamaño 3x3x3 fueron utilizados en todas las capas convolucionales. Todas las capas de pooling usan filtros de tamaño 1x2x2. Cada capa convolucional está seguida de un operador ReLU. Después las 256 características extraídas del proceso fueron aplanadas en un vector y este fue usado como el valor de entrada para las tres capas totalmente conectadas. Para limitar el riesgo del overfitting se utilizó un dropout de 0.5 en 2 de las capas totalmente conectadas. Finalmente se utilizó la operación soft-max en la última capa totalmente conectada para la clasificación.

## 4.4 Transformada de Hough circular

Tomando en cuenta la característica principal de SCPR de ser circular, una forma para reducir el tiempo para la localización y clasificación de estos se propone el uso del algoritmo de la transformada de Hough circular, que es este a su vez una especialización del algoritmo de la transformada de Hough, en específico el incorporado en la función `imfindcircles` en el programa `MatLab`[?].

La transformada de Hough es una técnica de extracción de características para imágenes, basada en la búsqueda de muestras imperfectas de ciertas clases de formas geométricas a través de un sistema de votos. El sistema de votos se lleva a cabo en un espacio de parámetros, del cual los objetos candidatos como un máximo local en un llamado espacio de acumulación que es construido de manera explícita a través del algoritmo de la transformada de Hough.

En un espacio bidimensional, un círculo puede ser descrito como:

$$(x - a)^2 + (y - b)^2 = r^2$$



donde  $(a,b)$  es el centro del círculo, y  $r$  el radio de este, si se fija el punto  $(x,y)$  el espacio de parametros sería tridimensional formado por  $(a,b,r)$ . Todos los parametros que satisfagan  $(x,y)$  se encuentran en un cono invertido de ángulo agudo cuyo vértice se encuentra en  $(x,y,0)$ . En un espacio tridimensional los parametros del círculo buscados pueden ser identificados como la intersección de varias superficies cónicas que son definidas de varios puntos en el círculo en el plano bidimensional. Este proceso puede dividirse en dos partes, la primera parte es fijar el radio y encontrar el centro óptimo de los círculos bidimensionales. La segunda parte es encontrar el radio óptimo en un espacio de parametros unidimensional.

El algoritmo de la transformada de Hough circular en la función `imfindcircles` en el programa MatLab

1. Calculo del espacio de acumulación

Los píxeles en primer plano de un gradiente mayor en comparación a los demás son designados como píxeles candidatos para emitir votos en el espacio de acumulación. Los píxeles candidatos votan en un patrón circular de radio fijo.

2. Estimación del centro

Los votos de los píxeles candidatos se tienden a acumular en el espacio de acumulación en el píxel que corresponde al centro

3. Estimación del radio

Si el mismo espacio de acumulación se ocupa para calcular mas de un radio, los radios de los círculos detectados tienen que estimarse en un proceso aparte.

La función `imfindcircles` dispone de 2 algoritmos para la localización de círculos en las imágenes: "Phase-Coding" y "Two-Stages", los cuales tienen sus características particulares pero también tienen en común:

- El uso de un espacio de acumulación bidimensional.  
La forma clásica de la transformada de Hough requiere el uso de un espacio de acumulación tridimensional para guardar los votos de múltiples radios, lo que resulta en grandes volúmenes de información almacenada y tiempos de computo. Tanto el algoritmo Phase-Coding y Two-Stages manejan el uso de un espacio de acumulación bidimensional el cual resuelve este problema, aunque viene con el problema inherente de que la estimación del radio se hace aparte, pero sigue siendo aun así la carga computacional menor.
- El uso de píxeles frontera La complejidad y el tiempo para realizar el problema está dado por la cantidad de píxeles candidatos asignados. Para limitar el número de píxeles candidatos, un criterio sobre el gradiente del píxel con respecto a la imagen se utiliza para desestimar todos los píxeles que no sobrepasen ese criterio.
- Uso de la información de la orientación de la frontera  
Otra herramienta para optimizar el desempeño de los píxeles candidatos es limitar la cantidad de votos que pueden emitir, usando la información direccional del gradiente, se puede tomar solo una región circular que sea congruente con esa dirección del gradiente.

## References

- [1] Xiao Xiang Zhu, Devis Tuia, Lichao Mou, Gui-Song Xia, Liangpei Zhang, Feng Xu, Friedrich Fraundorfer *Deep Learning in Remote Sensing: A Review*. IEEE Geoscience and Remote Sensing Magazine vol. 5, no. 4, pp 8-36 ,2017

- [2] P. Ghamisi, Y. Chen, and X. Zhu. *A self-improving convolution neural network for the classification of hyperspectral data*. IEEE Geoscience and Remote Sensing Letters, vol. 13, no. 10, pp. 1537–1541, 2016.
- [3] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, *Deep feature extraction and classification of hyperspectral images based on convolutional neural networks* IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 10, pp. 6232–6251, 2016.
- [4] Y. Li, H. Zhang, and Q. Shen. *Spectral-spatial classification of hyperspectral imagery with 3d convolutional neural network*. Remote Sensing, vol. 18, no. 7, pp. 1527–1554, 2006.
- [5] Chenxiao Zhang , Peng Yue , Liping Di and Zhaoyan Wu. *Automatic Identification of Center Pivot Irrigation Systems from Landsat Images Using Convolutional Neural Networks* Agriculture 2018, 8(10), 147
- [6] Hou, W., Wang, J., Xu, X., Reid, J. S., y Han, D. *An algorithm for hyperspectral remote sensing of aerosols: 1. Development of theoretical framework*. Journal of Quantitative Spectroscopy & Radiative Transfer, 178, 400–415.
- [7] Earth Explorer provisto por U.S. Geological Survey. Disponible online: <https://earthexplorer.usgs.gov>
- [8] Ciresan, D.C.; Meier, U.; Gambardella, L.M.; Schmidhuber, J. *Deep big simple neural nets excel on handwritten digit recognition*. Neural Comput. 2010, 22, 3207–3220.
- [9] Yu, X.; Wu, X.; Luo, C.; Ren, P. *Deep learning in remote sensing scene classification: A data augmentation enhanced convolutional neural network framework*. GISci. Remote Sens. 2017, 54, 741–758.
- [10] D. Tran, L. Bourdev, R. Fergus, L. Torresani, y M. Paluri, *Learning spatiotemporal features with 3d convolutional networks* IEEE International Geoscience and Remote Sensing Symposium, 2017
- [11] Documentacion de la funcion imfindcircles provista por la pagina oficial de Matlab. Disponible online:<https://www.mathworks.com/help/images/ref/imfindcircles.html>