

CAPITULO 6

Aplicación del razonamiento evidencial dinámico

6.1. Introducción

El método RED ha sido aplicado a la clasificación de cultivos en imágenes hiperespectrales de la región de Nebraska (USA). Esta región es rica en la producción de cereales, y sus cultivos más característicos son el maíz, la soja, el sorgo, y la alfalfa.

La empresa que proveyó las imágenes tanto multiespectrales como hiperespectrales, realiza estudios analizando el "stress" de los cultivos y se cuenta con los datos de referencia que los mismos utilizan. En la **Figura 6.1**, puede observarse parte de una de las imágenes utilizadas para la clasificación.

Para aplicar el clasificador se trabajó de la siguiente manera:

Se tomaron como entrada al clasificador las 10 bandas (b_i) espectrales que componen las imágenes hiperespectrales usadas. A las mismas se las utilizó para calcular la diferencia normalizada entre bandas consecutivas, es decir $(b_i - b_{i+1}) / (b_i + b_{i+1})$ para $i=1, \dots, 9$. El hecho de utilizar las diferencias normalizadas surge de la idea de acotar los valores dentro de un rango de manera que no exista tanta dispersión y el método no se vea influenciado por valores individuales, esto permite evitar en cierta manera el proceso de transformación de bin (utilizado para proveer una distribución del rango de valores - Ver **Capítulo 4**). Además se utilizan 4 rasgos de textura calculados sobre una de las bandas (760 nanómetros): la media, el "skewness" (medida de simetría), el rango y la desviación estándar, que son estadísticas de primer orden (ver **Capítulo 3 Sección 4**), y permiten incorporar información espacial al sistema y de este modo no tener sólo información espectral.

Esta decisión también está sustentada en el conocimiento que se tiene acerca de que los diferentes materiales de la superficie terrestre poseen una respuesta espectral diferente y las diferencias normalizadas permiten captar la información de los picos y valles que tienen dichas firmas espectrales (en particular las de la vegetación sana). Cuando la curva da un salto de un valor pequeño a uno mayor se obtendrán valores negativos y el valor absoluto de esta diferencia permitirá evaluar la intensidad de dicho salto (**Figura 6.2**). Para los distintos cultivos, si bien todos tienen una curva espectral similar, se encuentran diferencias que pueden detectarse cuando se tiene el detalle de una imagen hiperespectral. Se sabe que los tipos de cultivo específicos pueden ser identificados en base a los patrones de respuesta espectral y la textura de la imagen. El agregar los rasgos espaciales permite tener nuevas variables de discriminación para la clasificación. Las distintas formas de plantación, la altura, la forma de las hojas de los diferentes cultivos influyen sobre las características de textura de una imagen digital [**Jen96**][**Ped91**].

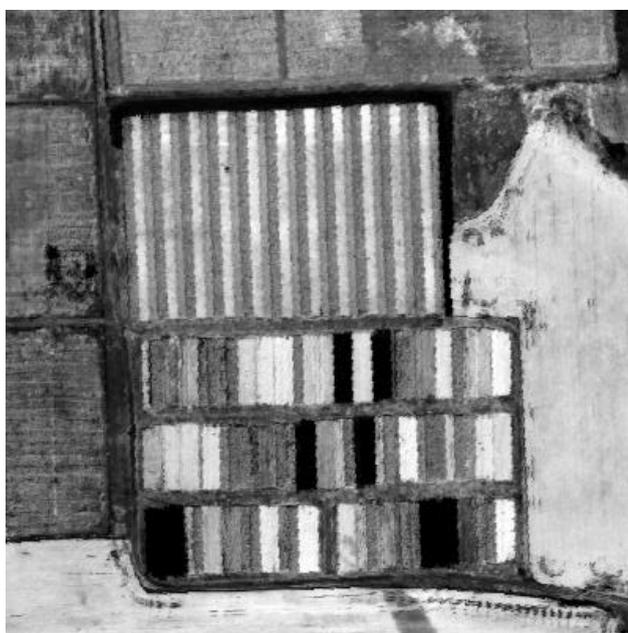


Figura 6.1 – Subimagen de la imagen 1m0809corn9.pix utilizada para la clasificación

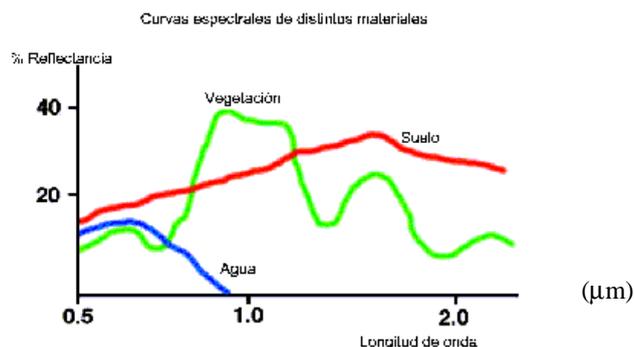


Figura 6.2 – Comparación de curvas espectrales de distintos materiales. Se muestra la reflectancia que caracteriza a estos materiales en las distintas bandas espectrales. Se puede observar, por ejemplo, que mientras que la vegetación tiene su máxima reflectancia en el cercano infrarrojo, el agua la tiene en el espectro visible. El eje x representa la longitud de onda medida en micrómetros

Es importante destacar en este punto, que la idea de esta tesis **no se centró en la búsqueda de las fuentes de información que resultaran óptimas para el método**. Es probable que exista un conjunto de fuentes que provean una mayor separabilidad para las clases en cuestión, y tal vez dicho conjunto sea aún menor en cardinalidad que el actual. En este trabajo **se puso como objetivo la prueba y análisis del método propuesto y su funcionamiento**, aún en casos donde exista una superposición entre las clases en un espacio N-dimensional de rasgos (de hecho, existe cierto solapamiento entre la clase sorgo y maíz, y aún así se obtienen los resultados esperados). Sin embargo, es cierto que previo a la elección de este conjunto de fuentes se estudiaron otros posibles con los cuales no se obtuvieron tan buenos resultados como con el conjunto seleccionado. Se probó trabajar con cada una de las bandas como una fuente de información, y con diferentes rasgos de textura.

La base de conocimiento inicial para el clasificador se formó a partir de un conjunto de muestras cuya pertenencia a una clase es conocida. Como muestras, por un lado se utilizaron píxeles de los que se extrajo su valor de intensidad en cada una de las bandas, y luego a partir de los mismos se calcularon las diferencias normalizadas. Por otro lado, se calcularon los valores de textura para cada una de dichas muestras, utilizando una ventana de 3x3 sobre una de las bandas espectrales (760 nm).

Se quiere resaltar la ventaja de poder incorporar información al sistema de *diferente naturaleza* como son los rasgos espaciales por un lado, y los espectrales por otro. Cada uno de estos rasgos se considera una fuente de información diferente para el razonamiento evidencial y el razonamiento evidencial dinámico.

En la **Figura 6.3**, se pueden observar las curvas espectrales de los conjuntos de muestras utilizados para el entrenamiento inicial.

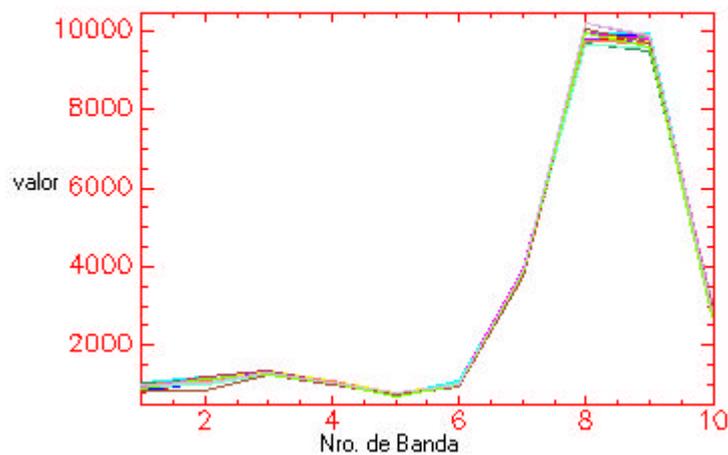


Figura 6.3 a – Muestras de la clase soja

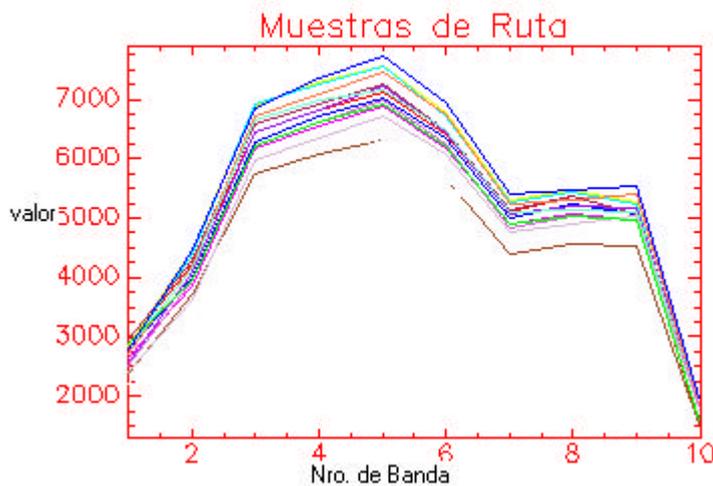


Figura 6.3 b – Muestras de la clase ruta



Figura 6.3 c – Muestras de la clase sorgo



Figura 6.3 d – Muestras de la clase maíz

Después de la etapa de entrenamiento inicial se prosiguió con la etapa de aprendizaje del clasificador con un conjunto de muestras de entrenamiento para las diferentes clases involucradas.

El primer conjunto de muestras de test, arrojó resultados que indicaban la necesidad de tener mayor cantidad de información para poder clasificar correctamente.

Por lo tanto, es necesario realizar el proceso de aprendizaje. Consecuentemente, se incorporaron las muestras a la base de conocimiento, y se calcularon los soportes para todo el conjunto nuevamente. Así, se procedió con cada uno de los conjuntos de test en este proceso de aprendizaje, y automáticamente, de acuerdo a los resultados obtenidos, el clasificador determinó si era necesario incorporar la información o no. La decisión del clasificador de aprender del conjunto de test o no hacerlo, se basa en el valor del indicador Khat. Sólo se aprende si el valor de Khat es inferior a un α considerado como el **parámetro de precisión**. Para este caso, se tomó un α igual a 0.90 equivalente a tener el 90% de precisión más que si se hiciera una clasificación al azar. De esta manera, el sistema sólo incorpora las muestras cuando se tiene una precisión por debajo de la indicada por α . Si Khat arroja valores inferiores a 0.90, no se ha alcanzado la precisión pretendida para el clasificador respecto de las muestras de entrenamiento. En la siguiente sección se analiza más detalladamente la influencia del parámetro de precisión.

El proceso de aprendizaje se repite presentando los mismos conjuntos de muestras hasta que no se produzcan modificaciones en la base de conocimiento. En la primer y segunda presentación del conjunto de muestras se utilizó como parámetro de precisión $\alpha=0.90$. No fue necesaria una tercera presentación ya que todos los conjuntos fueron clasificados con un valor de Khat superior a α .

Análisis del parámetro de precisión

El parámetro de precisión permite controlar el proceso de aprendizaje de forma tal de regular la incorporación o no de las muestras de entrenamiento al sistema. De acuerdo a los valores de Khat que arroja el clasificador, se puede ver cómo está funcionando el mismo para las muestras que se le presentan y si corresponde o no continuar aprendiendo. Se sabe que cuando Khat tiene un valor de 1 significa que el clasificador evita en un 100% los errores que se tendrían si se utilizara un clasificador completamente aleatorio. Cuando los valores de α se aproximan a 1, existe una mayor incorporación o aprendizaje en el sistema, es decir se agregan más cantidad de muestras a la base de conocimiento, esto conlleva por un lado a un

aumento en la precisión del clasificador y por otro, a un mayor tiempo de respuesta en posteriores clasificaciones. Por tal razón, se busca un equilibrio entre la precisión y el tiempo de respuesta esperados para la aplicación. Para esta clasificación se probaron diferentes valores para el parámetro de precisión, encontrando la mejor respuesta con el valor ya sugerido.

6.2. Resultados obtenidos.

6.2.1. Análisis del proceso de aprendizaje

En esta sección la intención es mostrar cómo funciona el proceso de aprendizaje para la aplicación en cuestión. Se presenta la forma en que el clasificador va logrando el aprendizaje a través de la incorporación de nueva evidencia.

Se muestran a continuación algunos resultados obtenidos en la aplicación del método RED en las imágenes descritas. Los resultados se evalúan utilizando la matriz de confusión, la precisión general y el indicador Khat.

Se seleccionaron conjuntos de entrenamiento diferentes de los utilizados para el entrenamiento inicial. Se realizó la etapa de aprendizaje del clasificador presentando las muestras de entrenamiento para cada clase en la forma descrita en el **Capítulo 3**. En una segunda pasada de los conjuntos de entrenamientos el valor de Khat para todos los casos resultó ser superior a α , indicando una alta precisión para dichos conjuntos.

Tabla 6.1 - Conjunto 1 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	74	6	5	0	0	85
Maíz	1	19	22	0	0	42
Sorgo	0	7	56	0	0	63
Ruta	9	3	0	37	0	49
Desc.	0	0	0	0	0	0
Total	84	35	83	37	0	239

Precisión general: 77%

Precisión por clase: soja (87%), maíz (45%), sorgo (88%), ruta (71%)

Khat: 0.69

El valor de Khat no logra superar el valor α (0.90) establecido para la aplicación, con lo cual se decide agregar la evidencia dada por el conjunto de muestras actual. Se incorpora el conjunto de muestras a la base de conocimiento del clasificador y se calculan nuevamente los soportes para cada fuente y clase. Luego, se evalúa el clasificador con el conjunto de muestras 2, utilizando la base de conocimiento modificada por la incorporación de las nuevas muestras.

Tabla 6.2 - Conjunto 2 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	84	0	0	0	0	84
Maíz	0	22	8	0	0	30
Sorgo	0	2	54	0	0	56
Ruta	0	0	0	17	0	17
Desc.	0	0	0	0	0	0
Total	84	24	62	17	0	187

Precisión general: 95 %

Precisión por clase: soja (100%), maíz (73%), sorgo (96%), ruta (100%)

Khat: 0.92

Para este caso, no se incorporan los datos del conjunto de muestras 2, ya que el clasificador arrojó un valor de Khat igual a 0.92 (92%), cercano al valor máximo que puede tomar el indicador. Esto implica que el clasificador cuenta con la evidencia necesaria para arrojar el resultado adecuado para la clasificación del conjunto de muestras actuales, y por tanto no modifica su conocimiento.

Tabla 6.3 - Conjunto 3 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	72	2	3	0	0	77
Maíz	0	16	8	0	0	24
Sorgo	0	2	88	0	0	90
Ruta	0	0	0	30	0	30
Desc.	0	0	0	0	0	0
Total	72	20	94	30	0	221

Precisión general: 93%

Precisión por clase: soja (93%), maíz (67%), sorgo (95%), ruta (100%)

Khat: 0.89

Se vuelven a incorporar los datos, en este caso del conjunto de test3, ya que el clasificador arrojó un valor de Khat inferior a 0.90.

Tabla 6.4 - Conjunto 4 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	89	1	0	0	0	90
Maíz	0	10	0	0	0	10
Sorgo	0	4	92	0	0	96
Ruta	0	2	2	50	0	54
Desc.	0	0	0	0	0	0
Total	89	17	94	50	0	250

Precisión general: 96 %

Precisión por clase: soja (99%), maíz (100%), sorgo (97%), ruta (92%)

Khat: 0.94

El clasificador no incorpora las muestras, dado el valor de Khat indica un error inferior al 10%.

Tabla 6.5 - Conjunto 5 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	64	0	0	0	0	64
Maíz	2	24	10	0	0	36
Sorgo	0	3	97	0	0	100
Ruta	0	0	0	30	0	30
Desc.	0	0	0	0	0	0
Total	66	27	107	30	0	230

Precisión general: 93%

Precisión por clase: soja (100%), maíz (67%), sorgo (97%), ruta (100%)

Khat: 0.90

El clasificador no incorpora esta información, debido a los buenos resultados obtenidos.

Tabla 6.6 - Conjunto 6 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	88	1	1	0	0	90
Maíz	2	15	2	1	0	20
Sorgo	0	5	125	0	0	130
Ruta	0	0	0	30	0	30
Desc.	0	0	0	0	0	0
Total	90	21	128	31	0	270

Precisión general: 95%

Precisión por clase: soja (98%), maíz (75%), sorgo (95%), ruta (100%)

Khat: 0.93

Tabla 6.7 - Conjunto 7 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	64	0	0	0	0	64
Maíz	1	38	17	0	0	56
Sorgo	1	4	85	0	0	90
Ruta	0	0	5	107	0	112
Desc.	0	0	0	0	0	0
Total	66	42	107	107	0	322

Precisión general: 91%

Precisión por clase: soja (100%), maíz (69%), sorgo (94%), ruta (95%)

Khat: 0.88

Las muestras del conjunto 6 no se incorporan y se incorporan las muestras del conjunto 7.

El porcentaje promedio para la primera pasada resultó ser de 87 %.

Luego de realizar la primera pasada de los conjuntos de entrenamiento, se volvieron a presentar los mismos al clasificador y no fue necesario realizar una tercer pasada ya que no hubo modificación de la base de datos de conocimiento en la pasada anterior (es decir que se obtuvo un valor superior o igual a **a** para todos los conjuntos presentados).

Se muestran a continuación los resultados obtenidos para los conjuntos de test en la segunda pasada:

Tabla 6.8 - Conjunto 1 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	85	0	0	0	0	85
Maíz	0	42	0	0	0	42
Sorgo	0	0	63	0	0	63
Ruta	0	0	0	49	0	49
Desc.	0	0	0	0	0	0
Total	85	42	63	49	0	249

Precisión general: 100%

Precisión por clase: soja (100%), maíz (100%), sorgo (100%), ruta (100%)

Khat: 1.00

Tabla 6.9 - Conjunto 2 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	84	0	0	0	0	84
Maíz	0	29	1	0	0	30
Sorgo	0	0	56	0	0	56
Ruta	0	0	0	17	0	17
Desc.	0	0	0	0	0	0
Total	84	29	57	17	0	187

Precisión general: 99%

Precisión por clase: soja (100%), maíz (97%), sorgo (100%), ruta (100%)

Khat: 0.99

Tabla 6.10 - Conjunto 3 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	77	0	0	0	0	77
Maíz	0	23	1	0	0	24
Sorgo	0	0	90	0	0	90
Ruta	0	0	0	30	0	30
Desc.	0	0	0	0	0	0
Total	77	24	90	30	0	221

Precisión general: 100%

Precisión por clase: soja (100%), maíz (96%), sorgo (100%), ruta (100%)

Khat: 0.99

Tabla 6.11 - Conjunto 4 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	89	1	0	0	0	90
Maíz	0	8	2	0	0	10
Sorgo	0	7	89	0	0	96
Ruta	1	1	1	51	0	54
Desc.	0	0	0	0	0	0
Total	90	17	92	51	0	250

Precisión general: 95%

Precisión por clase: soja (99%), maíz (80%), sorgo (93%), ruta (95%)

Khat: 0.92

Tabla 6.12 - Conjunto 5 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	64	0	0	0	0	64
Maíz	0	32	4	0	0	36
Sorgo	0	5	95	0	0	100
Ruta	0	0	0	30	0	30
Desc.	0	0	0	0	0	0
Total	64	37	99	30	0	230

Precisión general: 96%

Precisión por clase: soja (100%), maíz (95%), sorgo (96%), ruta (100%)

Khat: 0.94

Tabla 6.13 - Conjunto 6 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	88	1	1	0	0	90
Maíz	2	16	2	0	0	20
Sorgo	0	5	125	0	0	130
Ruta	0	0	0	30	0	30
Desc.	0	0	0	0	0	0
Total	91	21	128	30	0	270

Precisión general: 96%

Precisión por clase: soja (98%), maíz (80%), sorgo (96%), ruta (100%)

Khat: 0.94

Tabla 6.14 - Conjunto 7 - Matriz de confusión

	Soja	Maíz	Sorgo	Ruta	Desc.	Total
Soja	64	0	0	0	0	64
Maíz	0	56	0	0	0	56
Sorgo	0	0	90	0	0	90
Ruta	0	0	0	112	0	112
Desc.	0	0	0	0	0	0
Total	64	56	90	112	0	322

Precisión general: 100%

Precisión por clase: soja (100%), maíz (100%), sorgo (100%), ruta (100%)

Khat: 1.00

Como se explicó anteriormente, dado que no hubo modificación de la base de conocimiento, no se continúa con el proceso de aprendizaje.

A continuación se resaltan algunos puntos de este proceso:

- ✓ El porcentaje promedio de valor de Khat obtenido en la primera pasada es de 87.85 %, mientras que el obtenido en la segunda pasada es de 96.85 %. Esto indica una mejora respecto de la respuesta del clasificador luego del proceso de aprendizaje.
- ✓ Respecto de los valores de clasificación por cada clase, analizamos como se modifican a lo largo del proceso de entrenamiento:

Tabla 6.15 – Comparación entre los valores promedios obtenidos en la 1º y 2º pasada

Clase	1ra. Pasada Promedio	2da. Pasada Promedio
Soja	96.71 %	99.57 %
Maíz	70.85 %	92.57 %
Sorgo	94.57 %	97.85 %
Ruta	94 %	99.28 %

Esta tabla permite afirmar que los promedios de respuesta para cada clase son superiores en la segunda pasada del proceso de aprendizaje, lo cual demuestra una mejora introducida por este proceso en el comportamiento del clasificador para el conjunto de muestras de entrenamiento.

- ✓ Estos resultados dan la posibilidad de afirmar el buen comportamiento del clasificador RED para las muestras de entrenamiento, pero todavía falta probar al mismo para conjunto de test diferentes de las muestras de entrenamiento, a fin de tener una evaluación real del clasificador.

6.2.2 Evaluación del Clasificador y de su regla de decisión

La sección anterior estuvo dedicada al estudio del proceso de aprendizaje que utiliza RED. En esta sección se analiza, por un lado cómo se comporta el clasificador utilizando conjuntos de test diferentes a los utilizados como entrenamiento, y por otro lado cómo influye la regla de decisión incorporada en la clasificación.

Para poder entender en qué medida influye la regla de decisión se compara con otras alternativas como la de seleccionar el máximo soporte. De esta forma se evalúan y contraponen los resultados de cada una. Además se investiga cómo funcionan las dos etapas de rechazo propuestas por RED, especialmente se tiene en cuenta cómo afectan al funcionamiento del clasificador.

Con el objetivo de medir la precisión del clasificador se seleccionaron regiones de test, para las diferentes clases con las que se trabaja, de 4 imágenes hiperespectrales diferentes (para más referencias sobre el material de trabajo ver **Capítulo 3**). Es importante destacar que dichas regiones no coinciden con las utilizadas para el entrenamiento. Fueron elegidas aleatoriamente y dependiendo de los datos de referencia con los que se cuenta (por ejemplo, no es posible seleccionar igual cantidad de muestras para todas las clases dado que no todas las imágenes presentan los tres tipos de cultivos).

6.2.2.1 Análisis de la precisión de RED para diferentes conjuntos de test

A continuación se presentan los resultados obtenidos de la aplicación de RED a dichos conjuntos de test:

Tabla 6.16 - Clase Maíz

Clase Maíz	Porcentaje asignado a Clase:				
	Soja	Maíz	Sorgo	Ruta	Desc.
Región 1. Total: 49 píxeles	0 %	86 %	14 %	0 %	0 %
Región 2. Total: 28 píxeles	0 %	89 %	10 %	0 %	0 %
Región 3. Total: 64 píxeles	1.5 %	81 %	17 %	0 %	0 %
Región 4. Total: 30 píxeles	0 %	90 %	10 %	0 %	0 %
Región 5. Total: 40 píxeles	0 %	90 %	10 %	0 %	0 %
Región 6. Total: 100 píxeles	2 %	88 %	10 %	0 %	0 %
Región 7. Total: 84 píxeles	2 %	89 %	8 %	0 %	0 %
Región 8. Total: 10 píxeles	0 %	100 %	0 %	0 %	0 %
Región 9. Total: 12 píxeles	0 %	100 %	0 %	0 %	0 %
Región 10. Total: 25 píxeles	4 %	80 %	16 %	0 %	0 %
Región 11. Total: 20 píxeles	0 %	95 %	5 %	0 %	0 %
Región 12. Total: 8 píxeles	0 %	87 %	12 %	0 %	0 %
Región 13. Total: 30 píxeles	0 %	96 %	3 %	0 %	0 %

Porcentaje promedio para la clase maíz: 90.07 %

Tabla 6.17 - Clase Ruta

Clase Ruta	Porcentaje asignado a Clase:				
	Soja	Maíz	Sorgo	Ruta	Desc.
Región 1. Total: 38 pixeles	0 %	0 %	0 %	100 %	0 %
Región 2. Total: 36 pixeles	0 %	0 %	0 %	100 %	0 %
Región 3. Total: 75 pixeles	1 %	0 %	0 %	98 %	0 %
Región 4. Total: 56 pixeles	0 %	0 %	0 %	100 %	0 %
Región 5. Total: 56 pixeles	0 %	0 %	0 %	100 %	0 %
Región 6. Total: 54 pixeles	0 %	0 %	0 %	100 %	0 %
Región 7. Total: 32 pixeles	0 %	0 %	0 %	100 %	0 %
Región 8. Total: 60 pixeles	0 %	0 %	0 %	100 %	0 %
Región 9. Total: 40 pixeles	0 %	0 %	0 %	100 %	0 %
Región 10. Total: 26 pixeles	0 %	0 %	0 %	100 %	0 %
Región 11. Total: 33 pixeles	0 %	0 %	0 %	100 %	0 %
Región 12. Total: 64 pixeles	0 %	0 %	0 %	100 %	0 %
Región 13. Total: 68 pixeles	0 %	0 %	0 %	100 %	0 %
Región 14. Total: 40 pixeles	0 %	0 %	0 %	100 %	0 %

Porcentaje promedio para la clase ruta: 99.85 %

Tabla 6.18 - Clase Soja

Clase Soja	Porcentaje asignado a Clase:				
	Soja	Maíz	Sorgo	Ruta	Desc.
Región 1. Total: 90 pixeles	98 %	1 %	0 %	0 %	0 %
Región 2. Total: 84 pixeles	98 %	1 %	0 %	0 %	0 %
Región 3. Total: 70 pixeles	100 %	0 %	0 %	0 %	0 %
Región 4. Total: 120 pixeles	98 %	1 %	0 %	0 %	0 %
Región 5. Total: 48 pixeles	100 %	0 %	0 %	0 %	0 %
Región 6. Total: 63 pixeles	100 %	0 %	0 %	0 %	0 %
Región 7. Total: 72 pixeles	97 %	2 %	0 %	0 %	0 %
Región 8. Total: 32 pixeles	90 %	6 %	3 %	0 %	0 %
Región 9. Total: 36 pixeles	94 %	2 %	2 %	0 %	0 %
Región 10. Total: 63 pixeles	89 %	9 %	1 %	0 %	0 %

Porcentaje promedio para la clase maíz: 96.4 %

Tabla 6.19 - Clase Sorgo

Clase Sorgo	Porcentaje asignado a Clase:				
	Soja	Maíz	Sorgo	Ruta	Desc.
Región 1. Total: 255 pixeles	0 %	2.7 %	97 %	0 %	0 %
Región 2. Total: 312 pixeles	0.6 %	5 %	94 %	0 %	0 %
Región 3. Total: 70 pixeles	1.4 %	0 %	98 %	0 %	0 %
Región 4. Total: 70 pixeles	0 %	7 %	93 %	0 %	0 %
Región 5. Total: 72 pixeles	1 %	19 %	79 %	0 %	0 %

Porcentaje promedio para la clase maíz: 92.2 %

Por lo que se puede observar en las tablas anteriores se tiene un porcentaje promedio superior o igual al 90% para todos los casos. Respecto de los porcentajes

individuales para cada región, se puede afirmar que todos ellos se encuentran por encima del 80%.

6.2.2.2 Análisis de la influencia de la regla de decisión propuesta por RED

Con el fin de examinar en qué medida influye la regla de decisión, se incluye un nuevo punto de evaluación que consiste en observar cuántos píxeles son rechazados por las dos etapas propuestas por RED. Para ello, se utilizan algunas de las regiones de entrenamiento y otras de test seleccionadas aleatoriamente, a continuación se presentan los valores obtenidos, en la última columna de las tablas 1-3 se muestra la cantidad de píxeles rechazados por ambigüedad que fueron luego clasificados correctamente por la regla propuesta por RED:

Tabla 6.20 a -

Conjunto de Entrenamiento: 1	Rechazo por falta de evidencia	Rechazo por ambigüedad	Cantidad de píxeles rechazados asignados a la clase correcta
Soja	0 píxeles	7 píxeles	7 píxeles (100 %)
Maíz	0 píxeles	18 píxeles	18 píxeles (100 %)
Sorgo	0 píxeles	6 píxeles	6 píxeles (100 %)
Ruta	0 píxeles	0 píxeles	0 píxeles

Tabla 6.20 b -

Conjunto de Entrenamiento: 6	Rechazo por falta de evidencia	Rechazo por ambigüedad	Cantidad de píxeles rechazados asignados a la clase correcta
Soja	0 píxeles	16 píxeles	14 píxeles (87.5 %)
Maíz	0 píxeles	17 píxeles	14 píxeles (82.3 %)
Sorgo	0 píxeles	27 píxeles	21 píxeles (78 %)
Ruta	0 píxeles	8 píxeles	8 píxeles (100 %)

Tabla 6.20 c -

Regiones de Test	Rechazo por falta de evidencia	Rechazo por ambigüedad	Cantidad de píxeles rechazados asignados a la clase correcta
Región 1 Soja (90 píxeles)	0 píxeles	25 píxeles	24 píxeles (96 %)
Región 1 Maíz (49 píxeles)	0 píxeles	42 píxeles	39 píxeles (92,8 %)
Región 1 Sorgo (255 píxeles)	0 píxeles	68 píxeles	60 píxeles (88.8 %)
Región 1 Ruta (38 píxeles)	0 píxeles	14 píxeles	14 píxeles (100 %)

Como se puede ver a través de las tablas anteriores (**Tablas 6.20 a-c**), un alto porcentaje de los pixeles rechazados en la etapa de “Rechazo por ambigüedad”, son clasificados correctamente utilizando el índice que combina cantidad de fuentes, los pesos de los pares fuente-clase y el soporte. Para los ejemplos mostrados no hay pixeles rechazados por la etapa de "Rechazo por falta de evidencia", aunque se ha visto en otras regiones que algunos pixeles son asignados a la clase desconocida por esta etapa. Se evaluó que a medida que se incrementa el valor del parámetro de cantidad de fuentes aportando evidencia a favor de la clase, aumenta la cantidad de pixeles asignados a la clase desconocida por esta etapa de rechazo. A continuación se muestra una tabla donde figuran los promedios de cantidad de fuentes para los pixeles de las diferentes regiones evaluadas.

Histogramas de Cantidad de Fuentes

Tabla 6.21 - Región de Entrenamiento 1

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Clases														
Soja	0	0	0	0	0	0	0	0	0	3	16	26	40	0
Maíz	0	0	0	0	0	1	1	0	1	0	0	12	27	0
Sorgo	0	0	0	0	0	0	0	0	0	1	9	21	32	0
Ruta	0	0	0	0	0	0	0	0	0	0	4	25	20	0

Tabla 6.22 - Conjunto de Entrenamiento 2

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Clases														
Soja	0	0	0	0	0	6	15	23	21	17	2	0	0	0
Maíz	0	0	0	5	1	11	7	5	1	0	0	0	0	0
Sorgo	0	0	0	0	0	1	5	14	19	13	4	0	0	0
Ruta	0	0	0	1	0	2	4	5	4	1	0	0	0	0

Tabla 6.23 - Conjunto de Entrenamiento 3

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Clases														
Soja	0	0	0	0	0	1	0	0	1	3	14	28	30	0
Maíz	0	0	0	0	0	2	0	1	0	0	2	8	11	0
Sorgo	0	0	0	0	0	0	0	0	1	7	18	32	32	0
Ruta	0	0	0	0	0	0	0	0	1	1	12	6	10	0

Tabla 6.24 - Conjunto de Entrenamiento 4

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Clases														
Soja	0	0	0	0	1	10	15	30	25	9	0	0	0	0
Maíz	0	0	0	0	1	4	1	2	2	0	0	0	0	0
Sorgo	0	0	0	0	0	1	16	20	30	18	9	2	0	0
Ruta	0	0	2	4	8	14	8	14	3	1	0	0	0	0

Tabla 6.25 - Conjunto de Entrenamiento 5

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Clases														
Soja	0	0	0	0	0	8	8	27	12	9	0	0	0	0
Maíz	0	0	0	1	3	7	9	10	6	0	0	0	0	0
Sorgo	0	0	0	0	1	5	11	34	34	14	1	0	0	0
Ruta	0	0	0	2	2	8	11	5	1	1	0	0	0	0

Tabla 6.26 - Conjunto de Entrenamiento 6

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Clases														
Soja	0	0	1	0	1	7	17	33	22	8	1	0	0	0
Maíz	0	0	0	3	9	1	3	4	0	0	0	0	0	0
Sorgo	0	0	0	0	0	2	13	35	38	36	6	0	0	0
Ruta	0	0	1	1	2	6	10	8	2	0	0	0	0	0

Tabla 6.27 - Conjunto de Entrenamiento 7

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Clases														
Soja	0	0	0	0	0	0	0	0	0	2	18	22	22	0
Maíz	0	0	0	0	1	2	2	0	0	1	2	19	29	0
Sorgo	0	0	0	0	0	0	1	0	0	2	12	45	30	0
Ruta	0	0	0	0	0	0	0	0	4	6	23	47	32	0

Haciendo una lectura rápida de la información condensada en estas tablas, se observa que para los conjuntos de entrenamiento se tiene que la cantidad de fuentes aportando evidencia se concentran entre 5 y 12, prácticamente no hay casos con sólo 2 fuentes aportando evidencia (se registran 5 casos), y no hay ningún caso donde sólo 1 fuente aporte evidencia. Por lo tanto, para los conjuntos de entrenamientos (esto es una vez llevado a cabo el proceso de aprendizaje), no aparecen casos donde se rechace por falta de evidencia. Esto parece tener lógica dado que precisamente se ha entrenado al clasificador a partir de estas muestras, y por tanto existe al menos un mínimo de evidencia para ellas.

Se verá ahora que ocurre con algunos de los conjuntos de test:

Tablas 6.28- Muestra el Histograma de cantidad de fuentes de diferentes regiones de test

Tabla 6.28 a

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Región 1 de Soja	0	0	0	1	3	4	7	30	23	17	5	0	0	0

Tabla 6.28 b -

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Región 1 de Sorgo	0	0	0	1	2	11	30	69	85	46	9	0	2	0
Región 2 de Sorgo	0	0	0	0	3	10	39	81	93	72	14	0	0	0
Región 2 de Sorgo	0	0	0	0	0	3	9	14	26	12	6	0	0	0

Tabla 6.28 c -

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Región 1 de Ruta	0	0	0	0	3	3	11	11	8	2	0	0	0	0
Región 2 de Ruta	0	0	0	1	0	5	6	12	7	5	0	0	0	0

Tabla 6.28 d -

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Región 6 de Maíz	0	0	3	11	19	23	30	11	3	0	0	0	0	0
Región 7 de Maíz	0	0	0	10	13	16	23	17	5	0	0	0	0	0

Para el caso de las regiones de test presentadas se puede observar la mayor cantidad concentrada entre 3 y 10 (es decir, las clases con mayor soporte presentan evidencia aportada por este número de fuentes principalmente)

Se muestran a continuación, algunos de los casos de región desconocida para el algoritmo, donde varía la distribución anterior y adquiere sentido el rechazo por falta de evidencia.

Tabla 6.29 – Regiones desconocidas por el algoritmo tomadas de la imagen “Corn”

Cant. Fuentes	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Región 1 Desconocida	0	14	61	44	13	0	0	0	0	0	0	0	0	0
Región 2 Desconocida	0	9	36	39	10	2	2	2	0	0	0	0	0	0

Es importante observar cómo se concentran ahora principalmente entre 1 y 4 fuentes aportando evidencia. Este estudio abre una nueva posibilidad de análisis para la etapa de "Rechazo por falta de Evidencia", donde en lugar de considerar la cantidad de

fuentes aportando evidencia, se podría evaluar la distribución de los histogramas anteriores, para dilucidar si se trata de una región perteneciente a una de las clases conocidas o no.

6.2.2.3 Análisis comparativo de RED

Se realiza en esta sección un análisis comparativo entre los resultados obtenidos utilizando RED (los mostrados anteriormente en la **Sección 6.2.2.1**) y los obtenidos, para los mismos conjuntos de test, con el algoritmo de Razonamiento evidencial utilizando como regla la del máximo soporte (ERMS). La base de conocimiento con la que se trabaja es la misma para ambos algoritmos. Como puede observarse en las **Figuras 6.3 – 6.6**, con RED se logran en la mayoría de los casos porcentajes de acierto mayores que para el algoritmo ERMS. Las tablas presentan los porcentajes de pixeles asignados por cada algoritmo a la clase correcta (es decir se sabe que la región a clasificar pertenece a dicha clase).

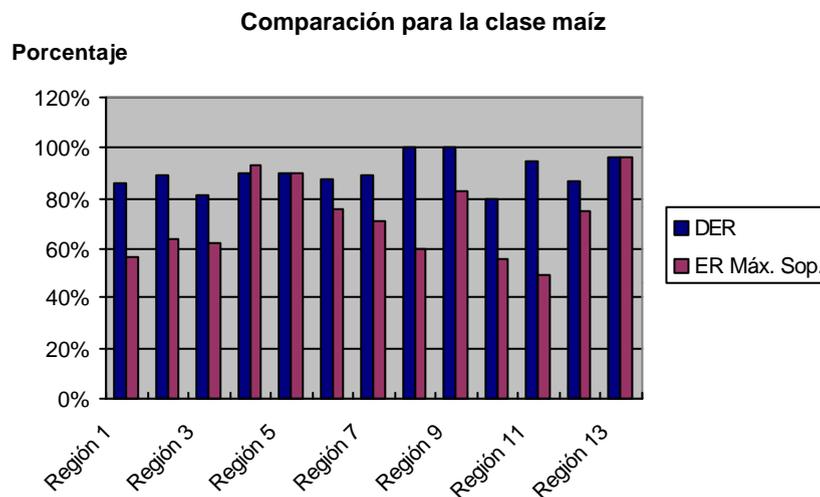


Figura 6.3 - Esta tabla presenta una comparación de los resultados de los algoritmos DER y ER con la regla de máximo soporte para la clase maíz

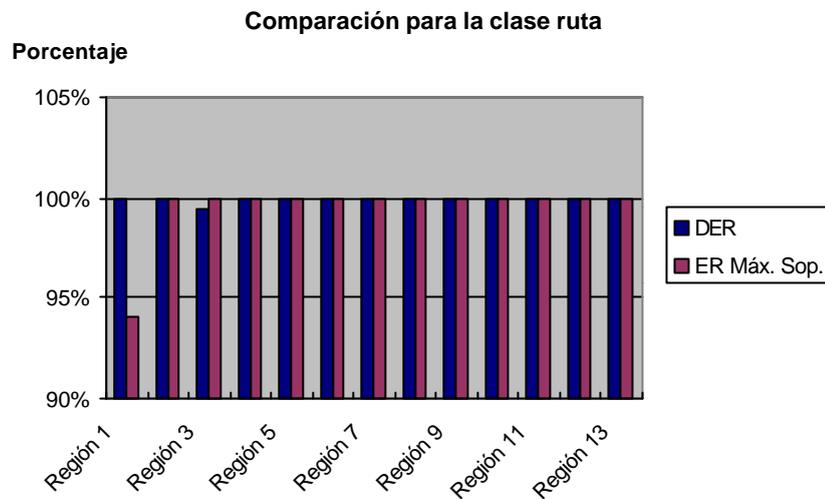


Tabla 6.31- Esta tabla presenta una comparación de los resultados de los algoritmos DER y ER con la regla de máximo soporte para la clase ruta

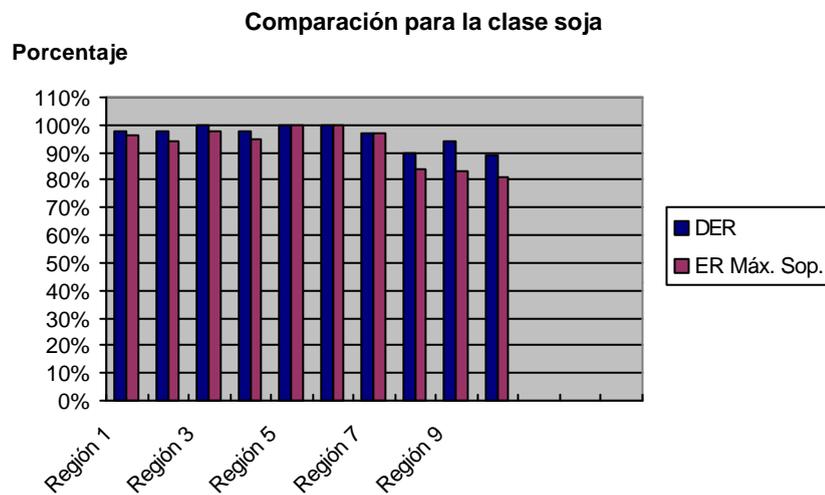


Tabla 6.32 - Esta tabla presenta una comparación de los resultados de los algoritmos DER y ER con la regla de máximo soporte para la clase soja

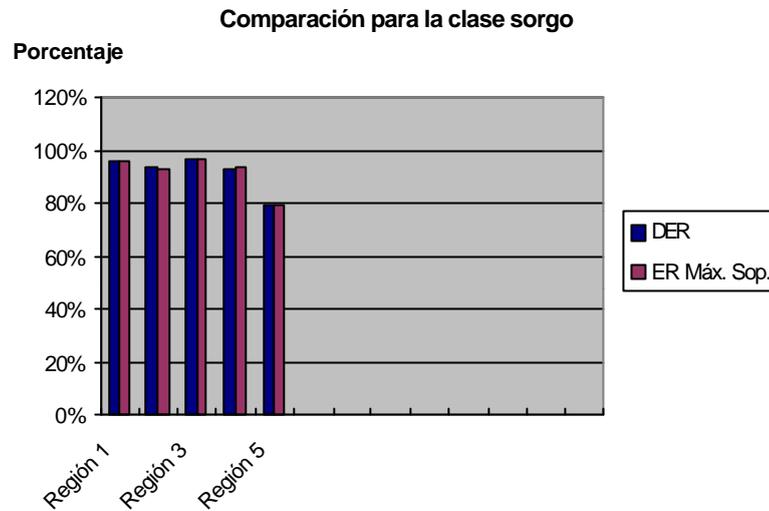


Tabla 6.33 - Esta tabla presenta una comparación de los resultados de los algoritmos DER y ER con la regla de máximo soporte para la clase sorgo

Clase Maíz

RED: 90.07 % ER: 71.76 %

Clase Ruta

RED: 99.85 % ER: 99.57 %

Clase Soja

RED: 96.4 % ER: 92.9 %

Clase Sorgo

RED: 91.8 % ER: 91.8 %

Se desprende de los resultados anteriormente expuestos que RED presenta un mayor o igual porcentaje promedio para todas las clases con las que se trabajó. En cuanto a los valores individuales de porcentaje para cada región clasificada, como se dijo anteriormente RED supera o iguala en la mayoría de los casos a los porcentajes obtenidos con ERMS.

Por otra parte, se estudió el algoritmo de razonamiento evidencial utilizando como regla de decisión la que elige la clase con mayor soporte más plausibilidad [Ped95 a]. Este se probó para las mismas regiones de test con las que se realizó la comparación anterior, y los porcentajes obtenidos resultaron similares a los de ER con la regla de soporte máximo.

Otras pruebas realizadas consistieron en evaluar el algoritmo RED pero utilizando igual regla de decisión, con la salvedad de asignar a la clase desconocida cuando se produce rechazo por ambigüedad. Se observó en esta experimentación que gran parte de los píxeles eran asignados a la clase desconocida. Esto puede verse en las tablas de la **Sección 6.2.2.3**, donde se presentan los píxeles por región que son rechazados por ambigüedad. Si se asignan los mismos a la clase desconocida, ésta se vería inundada de píxeles.

Para completar este análisis de resultados se presentan algunos valores obtenidos para regiones cuya clase es desconocida para el algoritmo. La mejor precisión obtenida aquí es la del algoritmo RED que asigna cuando encuentra rechazo por ambigüedad a la clase desconocida. En el caso de RED propiamente dicho, distribuye la asignación de píxeles entre las diferentes clases existentes, de igual forma se comporta ERMS, esto puede mejorarse teniendo en cuenta la distribución de la cantidad de fuentes aportando evidencia (histograma de cantidad de fuentes, ver **Sección 6.2.2.2**). También debajo se presentan casos donde se trabaja con muestras que pertenecieron al conjunto de entrenamiento, de forma tal de comparar cómo responden RED y ERMS ante las mismas. Se muestran a continuación, sólo algunos de los resultados obtenidos utilizando regiones desconocidas y otras de entrenamiento.

Región desconocida 1 de Imagen Bug 1. Se trata de una región de árboles de 121 píxeles.

Precisión por clase utilizando RED:

soja (17%), maíz (58%), sorgo (22%), ruta (0.8%), desconocida (0.8%)

Precisión por clase utilizando RED con asignación a la clase desconocida en caso de rechazo por ambigüedad:

soja (0 %), maíz (4.1%), sorgo (4.9%), ruta (0%), desconocida (90.9%)

Precisión por clase utilizando ERMS:

soja (7.43 %), maíz (56 %), sorgo (33%), ruta (2.4 %)

Región desconocida 2 de Imagen Bug 1. Se trata de una región de 49 píxeles.

Precisión por clase utilizando RED:

soja (24.4%), maíz (18.3%), sorgo (48.9%), ruta (8%)

Precisión por clase utilizando RED con asignación a la clase desconocida en caso de rechazo por ambigüedad:

soja (0 %), maíz (0%), sorgo (2%), ruta (0%), desconocida (97.9%)

Precisión por clase utilizando ERMS:

soja (24.4 %), maíz (18.3 %), sorgo (48.9%), ruta (8 %)

Conjunto de Entrenamiento 1

Precisión por clase utilizando RED:

soja (100%), maíz (100%), sorgo (100%), ruta (100%)

Precisión por clase utilizando RED con asignación a la clase desconocida en caso de rechazo por ambigüedad:

soja (91 %), maíz (57%, el resto asignado a la clase desconocida), sorgo (90%), ruta (100%)

Precisión por clase utilizando ERMS:

soja (100%), maíz (92%), sorgo (100%), ruta (100%)

Conjunto de Entrenamiento 2

Precisión por clase utilizando RED:

soja (100%), maíz (97%), sorgo (100%), ruta (100%)

Precisión por clase utilizando RED con asignación a la clase desconocida en caso de rechazo por ambigüedad:

soja (83 %), maíz (16%, el resto asignado a la clase desconocida), sorgo (87.5%), ruta (88%)

Precisión por clase utilizando ERMS:

soja (97%), maíz (83%), sorgo (96%), ruta (100%)

Conjunto de Entrenamiento 3

Precisión por clase utilizando RED:

Región soja (100%), Región maíz (96%), Región sorgo (100%), Región ruta (100%)

Precisión por clase utilizando RED con asignación a la clase desconocida en caso de rechazo por ambigüedad:

Región soja (87 %), Región maíz (37%, el resto asignado a la clase desconocida), Región sorgo (89%), Región ruta (100%)

Precisión por clase utilizando ERMS:

Región soja (98.7%), Región maíz (87.5%), Región sorgo (100%), Región ruta (100%)

6.3. Conclusiones del capítulo

En este capítulo se presentó una aplicación con la que se estudia el comportamiento del método RED propuesto en esta tesis. Se trabajó con imágenes de una región del estado de Nebraska, donde se analizan campos cultivados. Las imágenes son, como se explica en el **Capítulo 3**, obtenidas con un sensor *Casi*, y tienen tamaño aproximado de 55 MB cada una (para tener más información referirse al **Capítulo 3**).

En la primera sección se describe la aplicación, mientras que en las subsiguientes, se realiza el seguimiento de cómo se desarrolla el proceso de aprendizaje y cómo afecta la regla de decisión propuesta al momento de clasificar. Finalmente, se compara el método propuesto con el de razonamiento evidencial con regla del máximo soporte y la de máximo soporte más plausibilidad. También se analiza otra posible regla de decisión para RED. Se presentan los resultados obtenidos para un conjunto de regiones, surgidas de 4 imágenes hiperespectrales diferentes.

El método propuesto arroja los resultados esperados. Aún así, se debe evaluar el costo computacional que el mismo introduce respecto de la precisión necesaria para la aplicación. ◆