



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE EDAFOLOGÍA

**ANÁLISIS TEXTURAL Y DE COLOR DE IMÁGENES
DIGITALES DEL CULTIVO DE FRIJOL (*Phaseolus vulgaris*
L.) EN CONDICIONES DE DEFICIENCIA DE HIERRO Y
MANGANESO USANDO CLASIFICADORES CON REDES
NEURONALES ARTIFICIALES**

EDGAR GARCÍA CRUZ

T E S I S

**PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:**

DOCTOR EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MÉXICO

2018

CARTA DE CONSENTIMIENTO DE USO DE LOS DERECHOS DE AUTOR Y DE LAS REGALIAS COMERCIALES DE PRODUCTOS DE INVESTIGACION

En adición al beneficio ético, moral y académico que he obtenido durante mis estudios en el Colegio de Postgraduados, el que suscribe M.C. EDGAR GARCÍA CRUZ, Alumno (a) de esta Institución, estoy de acuerdo en ser participe de las regalías económicas y/o académicas, de procedencia nacional e internacional, que se deriven del trabajo de investigación que realicé en esta institución, bajo la dirección del Profesor DR. MANUEL SANDOVAL VILLA, por lo que otorgo los derechos de autor de mi tesis

ANÁLISIS TEXTURAL Y DE COLOR DE IMÁGENES DIGITALES DEL CULTIVO DE FRIJOL (Phaseolus vulgaris L.) EN CONDICIONES DE DEFICIENCIA DE HIERRO Y MANGANESO USANDO CLASIFICADORES CON REDES NEURONALES ARTIFICIALES

y de los productos de dicha investigación al Colegio de Postgraduados. Las patentes y secretos industriales que se puedan derivar serán registrados a nombre el colegio de Postgraduados y las regalías económicas que se deriven serán distribuidas entre la Institución, El Consejero o Director de Tesis y el que suscribe, de acuerdo a las negociaciones entre las tres partes, por ello me comprometo a no realizar ninguna acción que dañe el proceso de explotación comercial de dichos productos a favor de esta Institución.

Montecillo, Mpio. de Texcoco, Edo. de México, a 9 de abril de 2018

Firma del
Alumno (a)

DR. MANUEL SANDOVAL VILLA

Vo. Bo. del Consejero o Director de Tesis

La presente tesis titulada: ANÁLISIS TEXTURAL Y DE COLOR DE IMÁGENES DIGITALES DEL CULTIVO DE FRÍJOL (*Phaseolus vulgaris* L.) EN CONDICIONES DE DEFICIENCIA DE HIERRO Y MANGANESO USANDO CLASIFICADORES CON REDES NEURONALES ARTIFICIALES realizada por el alumno EDGAR GARCÍA CRUZ bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

DOCTOR EN CIENCIAS
EDAFOLOGÍA

CONSEJO PARTICULAR

CONSEJERO


DR. MANUEL SANDOVAL VILLA

ASESOR


DR. JOSÉ ALFREDO CARRILLO SALAZAR

ASESOR


DR. PAULINO PÉREZ RODRÍGUEZ

ASESOR


DR. JORGE D. ETCHEVERS BARRA

ASESORA


DRA. ANTONIA MACEDO CRUZ

Montecillo, Texcoco, Estado de México, abril de 2018

ANÁLISIS TEXTURAL Y DE COLOR DE IMÁGENES DIGITALES DEL CULTIVO DE FRIJOL (*Phaseolus vulgaris* L.) EN CONDICIONES DE DEFICIENCIA DE HIERRO Y MANGANESO USANDO CLASIFICADORES CON REDES NEURONALES ARTIFICIALES

Edgar García Cruz, D. en C.
Colegio de Postgraduados, 2018

RESUMEN

En la presente investigación se plantea el uso de métodos para la estimación del estatus nutrimental del hierro (Fe) y manganeso (Mn) en el cultivo de frijol (*Phaseolus vulgaris* L.). Se proponen como una alternativa futura a los métodos de análisis químico una vez que puedan ser validados, pero con las ventajas de ser una opción rápida, económica y de aplicación *in situ*. Tales características no solo constituyen las ventajas *per se*, sino que pueden ser utilizadas como una herramienta en la generación de nuevos modelos para la optimización de la nutrición de los elementos estudiados, la alimentación de sistemas expertos que unan los conocimientos de disciplinas separadas o la creación de herramientas digitales para el monitoreo del estatus nutrimental del cultivo. Bajo este contexto, este trabajo se divide en tres capítulos, cada uno de los cuáles aborda el uso de modelos con redes neuronales artificiales. En un primer capítulo, se estudia la relación entre el índice de verdor, medido con un SPAD 502 (Minolta, Corp.) y los días a lo largo de la fenología del cultivo con el contenido de Fe; se propone entonces un modelo con índice de correlación de Pearson de 0.71. Por otro lado, en el Capítulo II, se aborda también la ocurrencia de la deficiencia de Mn además de la deficiencia de Fe; experimentalmente, se generan condiciones para la expresión de estas deficiencias además de un exceso; posteriormente se identifican con clasificadores, mediante el uso de redes neuronales artificiales, los patrones de textura y color en dos momentos de muestreo. Se obtiene que, tanto las características de color como de textura pueden ser de utilidad para la identificación de las deficiencias de Fe y Mn en el cultivo de frijol, con un éxito en el rango de 70 a 80 %. Finalmente, cuando se considera el papel preponderante del nitrógeno (N) en la conformación de la molécula de clorofila y tomando en cuenta que, el Fe también es causa del mismo desorden fisiológico, se establece un experimento con el propósito de evaluar el efecto del N en la conformación de los patrones de textura y color en la identificación de las deficiencias de Fe. De forma adicional, se llevan a cabo cuatro muestreos con el fin de evaluar el comportamiento de dichos patrones. Se obtiene que, los clasificadores pueden diferenciar las diferentes condiciones nutrimentales de Fe interactuando con N en el orden de 76 %, considerando un porcentaje global de clasificaciones correctas en la etapa de prueba. Se requieren explorar otras características para mejorar el desempeño de los clasificadores, sin embargo, los resultados obtenidos en esta investigación se consideran convenientes.

Palabras clave: clorosis férrica; modelación; diagnóstico nutrimental; SPAD

**TEXTURAL AND COLOR ANALYSIS OF DIGITAL IMAGES OF THE BEAN CROP
(*Phaseolus vulgaris* L.) UNDER CONDITIONS OF IRON AND MANGANESE DEFICIENCY
USING CLASSIFIERS WITH ARTIFICIAL NEURAL NETWORKS**

**Edgar García Cruz, Ph. D.
Colegio de Postgraduados, 2018**

ABSTRACT

In the present investigation the use of methods for the estimation of the nutritional status of iron (Fe) and manganese (Mn) in the bean crop (*Phaseolus vulgaris* L.) is proposed. These options are proposed as a future alternative to chemical analysis methods once they can be validated; the advantages of these methods are a fast, economical and *in situ* application options. Such characteristics not only constitute the advantages *per se*, but can be used as a tool in the generation of new models for the optimization of the nutrition of the elements studied, the feed-back of expert systems that put together the knowledge of separate disciplines or the creation of digital tools for monitoring the nutritional status of the crop. Under this context, this work is divided into three chapters, each of which addresses the use of models with artificial neural networks. In a first chapter, the relationship between the green index, measured with a SPAD 502 (Minolta, Corp.) and the days along the phenology of the crop with the content of Fe is studied; a model with a Pearson correlation index of 0.71 is then obtained. On the other hand, in Chapter II, the occurrence of the deficiency of Mn in addition to the iron deficiency is also addressed; experimentally, conditions are generated for the expression of these deficiencies as well as an excess; subsequently, the texture and color patterns in two sampling moments are identified with classifiers, using artificial neural networks. It is obtained that, both the characteristics of color and texture can be useful for the identification of the deficiencies of Fe and Mn in the bean crop, with a success in the range of 70 to 80%. Finally, when considering the preponderant role of nitrogen (N) in the conformation of the chlorophyll molecule and taking into account that Fe is also the cause of the same physiological disorder, an experiment is established with the purpose of evaluating the effect of the N in the conformation of the texture and color patterns in the identification of the deficiencies of Fe. Additionally, four samples are carried out in order to evaluate the behavior of said patterns. It is obtained that the classifier can differentiate the different Fe nutritional conditions by interacting with N in the order of 76%, considering a global percentage of correct classifications in the test stage. It is necessary to explore other characteristics to improve the performance of the classifiers, however, the results obtained in this research are considered convenient.

Keywords: iron chlorosis; modeling; nutritional diagnosis; SPAD

D E D I C A T O R I A

*A ese Ser que jamás su espíritu ha sido quebrantado,
ni su alma doblegada. Te amo.*

*A aquellos rechazados e ignorados que convirtieron su rabia, furia
y frustración en arte, creación e inspiración.*

A G R A D E C I M I E N T O S

“A mitad del camino de la vida, en una selva oscura me encontraba porque mi ruta había extraviado.

¡Cuán dura cosa es decir cuál era esta salvaje selva, áspera y fuerte que me vuelve el temor al pensamiento!”

—Dante Alighieri

Esta investigación fue realizada gracias a los fondos concedidos por el CONSEJO NACIONAL DE CIENCIA Y TECNOLOGÍA, y con el apoyo académico del COLEGIO DE POSTGRADUADOS, institución a la cual me enorgullece haber pertenecido.

Mi agradecimiento pleno al Dr. Manuel Sandoval, Dr. Alfredo Carrillo, Dr. Paulino Pérez, Dr. Jorge Etchevers y Dra. Antonia Macedo por ser ustedes el centro de este proceso creativo e inspiracional.

La gratitud más intensa a mi familia; el principio y el origen. Gracias inmensas a mis amigos, mis compañeros en este viaje, y gracias a los que, su camino ha cambiado pero que han cumplido con un propósito de enseñanza y aprendizaje en algún ciclo.

Me es necesario agradecer también a aquellas personas que provocaron en mí el deseo mejorar en cualquiera de sus formas.

—Edgar García

C O N T E N I D O

RESUMEN	i
ABSTRACT	ii
LISTA DE FIGURAS	viii
LISTA DE CUADROS	x
INTRODUCCIÓN GENERAL	1
OBJETIVOS	12
Objetivo General	12
Objetivos Específicos.....	12
HIPÓTESIS	13
LITERATURA CITADA	14
CAPÍTULO I. MODELOS DE PREDICCIÓN Y CLASIFICACIÓN DE DEFICIENCIAS DE HIERRO EN EL CULTIVO DE FRIJOL (<i>Phaseolus vulgaris</i> L.) CON REDES NEURONALES BAYESIANAS REGULARIZADAS Y ÁRBOLES DE CLASIFICACIÓN	17
1.1 RESUMEN	17
1.2 SUMMARY	18
1.3 INTRODUCCIÓN	19
1.4 MATERIALES Y MÉTODOS	23
Establecimiento del experimento, trasplante y aplicación de tratamientos	23
Predicción de deficiencias de Fe usando redes neuronales bayesianas regularizadas (BRNN)	24
Predicción de deficiencias de Fe usando con árboles de clasificación.....	27
1.5 RESULTADOS Y DISCUSIÓN	28
Redes neuronales bayesianas regularizadas.....	28
Árboles de predicción y clasificación	31
1.6 CONCLUSIONES	36
1.7 LITERATURA CITADA	37
CAPÍTULO II. DETECCIÓN TEMPRANA Y CLASIFICACIÓN DE DEFICIENCIAS DE HIERRO Y MANGANESO EN FRIJOL	41

2.1 RESUMEN	41
2.2 SUMMARY	42
2.3 INTRODUCCIÓN	43
2.4 MATERIALES Y MÉTODOS	47
Siembra	47
Trasplante y tratamiento.....	47
Análisis de imágenes digitales y generación de los clasificadores con redes neuronales artificiales.....	48
Obtención de las imágenes digitales	48
Análisis de color y textura	49
Entrenamiento y prueba de escenarios de entradas	52
2.5 RESULTADOS Y DISCUSIÓN	54
2.6 CONCLUSIONES	62
2.7 LITERATURA CITADA	64
CAPÍTULO III. INTERACCIÓN DE DEFICIENCIAS DE HIERRO Y NITRÓGENO Y SU IDENTIFICACIÓN CON REDES NEURONALES ARTIFICIALES EN FRIJOL	67
3.1 RESUMEN	67
3.2 SUMMARY	68
3.3 INTRODUCCIÓN	69
3.4 MATERIALES Y MÉTODOS	71
Siembra	71
Trasplante y tratamiento.....	72
Análisis de imágenes digitales y generación de los clasificadores con redes neuronales artificiales.....	73
Obtención de las imágenes digitales	73
Análisis de color y textura	74
Entrenamiento y prueba de escenarios de entradas	76
Análisis de varianza	79
3.5 RESULTADOS Y DISCUSIÓN	79
3.6 CONCLUSIONES	92

3.7 LITERATURA CITADA	93
CONCLUSIONES GENERALES	95

LISTA DE FIGURAS

CAPÍTULO I

- Figura 1.1. Representación gráfica del modelo de red neuronal utilizada para $S = 3$ neuronas..... 25
- Figura 1.2. Representación gráfica tridimensional que muestra el comportamiento de las diferentes concentraciones de Fe en función del tiempo y el índice de verdor (valores SPAD) en el cultivo de frijol. Estos datos se usaron para el entrenamiento de la BRNN. Los puntos corresponden a las mediciones mientras que en la malla se reflejan los valores que corresponden a las predicciones obtenidas mediante la BRNN con $S = 3$ neuronas. 29
- Figura 1.3. Apariencia general de la deficiencia de Fe a los 18 ddt en el cultivo de frijol var. Negro Chapingo en dos fases distintas de desarrollo de la deficiencia pero en el mismo tratamiento (-Fe). 30
- Figura 1.4. Diagrama de cajas y bigotes para las correlaciones obtenidas mediante validación cruzada. Histograma de frecuencias de las diferentes correlaciones entre valores predichos contra valores observados obtenidas después de llevar a cabo el procedimiento de validación cruzada con 8 particiones aleatorias. 31
- Figura 1. 5. Predicción de Fe en el cultivo de frijol a partir del número de días después del trasplante (ddt) y los valores SPAD usando un (a) Diagrama de árbol y (b) Partición binaria del espacio..... 32

CAPÍTULO II

- Figura 2.1. Estrategia de entrenamiento tipo neuronal supervisada; esquema basado en el algoritmo de correlación en cascada propuesto por Fahlman y Lebiere (1990) para un clasificador con 16 variables de entrada y ocho variables de salida. Los números en los círculos indican las neuronas de la capa interna que se agregan una tras otra hasta llegar a 148. Los círculos vacíos indican conexiones congeladas, mientras que los círculos negros son entrenados repetidamente. 52

Figura 2.2. Cuarta hoja de frijol var. Cacahuate, recientemente madura, de una repetición de cada uno de los tratamientos indicados. Las hojas se recolectaron a los 63 días después de la siembra.....	59
Figura 2.3. Cuarta hoja de frijol var. Cacahuate, recientemente madura, de una repetición de cada uno de los tratamientos indicados. Las hojas se recolectaron a los 80 días después de la siembra.....	60

CAPÍTULO III

Figura 3.1. Estrategia de entrenamiento tipo neuronal supervisada; esquema basado en el algoritmo de correlación en cascada propuesto por Fahlman y Lebiere (1990) para un clasificador con 16 variables de entrada y nueve variables de salida para identificar interacciones de deficiencias de Fe y N en hojas de frijol var. Negro Chapingo a los 35 ddt. Los números en los círculos indican las neuronas de la capa interna que se agregan una tras otra hasta llegar a 126. Los círculos vacíos indican conexiones congeladas, mientras que los círculos negros son entrenados repetidamente.	78
Figura 3.2. Diagrama de procesos abreviados para un clasificador con 16 variables de entrada y nueve variables de salida para identificar interacciones de deficiencias de Fe y N en hojas de frijol var. Negro Chapingo para cada uno de los cuatro muestreos especificados.	78
Figura 3.3. Submuestras representativas del foliolo izquierdo de una repetición de cada uno de los tratamientos o clases en los cuatro momentos de muestreo (35, 56, 67y 78 ddt) de las hojas de frijol var. Negro Chapingo.....	82
Figura 3.4. Color promedio en el espacio RGB de cada uno de los tratamientos en los cuatro momentos de muestreo (35, 56, 67y 78 ddt) de las hojas de frijol var. Negro Chapingo.....	83

LISTA DE CUADROS

CAPÍTULO I

- Cuadro 1.1. Aciertos y errores de los valores observados contra valores modelados utilizando árboles de clasificación para la medición de especificidad/sensibilidad del modelo en entrenamiento..... 34
- Cuadro 1.2. Aciertos y errores para clase observada contra clase predicha utilizando árboles de clasificación para la medición de especificidad/sensibilidad del modelo en prueba. 34

CAPÍTULO II

- Cuadro 2.1. Tratamientos evaluados en la investigación con base en la solución Steiner (1984)..... 48
- Cuadro 2.2. Entrenamiento y prueba de escenarios de entradas para crear el clasificador con redes neuronales artificiales e identificar ocho clases de salida o niveles de fertilización con base en la textura y color de hojas de frijol en el primer muestreo (63 ddt): (1) Solución Steiner, Testigo; (2) Sol. Steiner +200 % Fe; (3) Sol. Steiner -50 % Fe; (4) Sol. Steiner –Fe; (5) Sol. Steiner –Fe y – Mn; (6) Sol. Steiner –Mn; (7) Sol. Steiner -50 % Mn y (8) Sol. Steiner +200 % Mn. 55
- Cuadro 2.3. Estadísticas promedio de 10 arreglos matriciales (tablas de contingencia) para el escenario de entradas: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la prueba, y ocho clases de salida: (1) Solución Steiner, Testigo; (2) Sol. Steiner +200% Fe; (3) Sol. Steiner -50% Fe; (4) Sol. Steiner – Fe; (5) Sol. Steiner –Fe y – Mn; (6) Sol. Steiner – Mn; (7) Sol. Steiner -50% Mn y (8) Sol. Steiner +200% Mn. Muestreo a los 63 ddt..... 55
- Cuadro 2.4. Entrenamiento y prueba de escenarios de entradas para crear el clasificador con redes neuronales artificiales e identificar ocho clases de salida o niveles de fertilización con base en la textura y color de hojas de frijol en el segundo muestreo (80 ddt): (1) Solución Steiner, Testigo; (2) Sol. Steiner +200 % Fe; (3) Sol. Steiner -50 % Fe; (4) Sol. Steiner –Fe; (5) Sol. Steiner –Fe y – Mn; (6) Sol. Steiner –Mn; (7) Sol. Steiner -50 % Mn y (8) Sol. Steiner +200 % Mn. 56

Cuadro 2.5. Estadísticas de 10 arreglos matriciales (tablas de contingencia) para la combinación de las variables de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH y ocho clases de salida: (1) Solución Steiner, Testigo; (2) Sol. Steiner +200% Fe; (3) Sol. Steiner -50% Fe; (4) Sol. Steiner – Fe; (5) Sol. Steiner –Fe y – Mn; (6) Sol. Steiner – Mn; (7) Sol. Steiner -50% Mn y (8) Sol. Steiner +200% Mn. Muestreo a los 80 ddt.....	57
--	----

CAPÍTULO III

Cuadro 3.1. Tratamientos aplicados con base en la solución Steiner (1984) a las plantas de frijol var. Negro Chapingo a los 15 dds.	72
--	----

Cuadro 3.2. Entrenamiento y prueba promedio de 10 particiones para crear el clasificador con redes neuronales artificiales y tres escenarios de entrada e identificar nueve clases de salida o niveles de fertilización con base en la textura y color de hojas de frijol: (1) Solución Steiner –N, –Fe; (2) Sol. Steiner –N, -50 % Fe; (3) Sol. Steiner –N; (4) Sol. Steiner –50 % N, –Fe; (5) Sol. Steiner –50 % N, –50 % Fe; (6) Sol. Steiner –50 % N; (7) Sol. Steiner –Fe; (8) Sol. Steiner –50 % Fe y (9) Solución Steiner, Testigo. Muestreo 1 (35 ddt), Muestreo 2 (56 ddt), Muestreo 3 (67 ddt) y Muestreo 4 (78 ddt).	80
--	----

Cuadro 3.3. Estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la prueba, y nueve clases de salida: (1) Solución Steiner –N, –Fe; (2) Sol. Steiner –N, -50 % Fe; (3) Sol. Steiner –N; (4) Sol. Steiner –50 % N, –Fe; (5) Sol. Steiner –50 % N, –50 % Fe; (6) Sol. Steiner –50 % N; (7) Sol. Steiner –Fe; (8) Sol. Steiner –50 % Fe y (9) Solución Steiner, Testigo. Muestreo 1 (35 ddt).	84
---	----

Cuadro 3.4. Estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la prueba, y nueve clases de salida: (1) Solución Steiner –N, –Fe; (2) Sol. Steiner –N, -50 % Fe; (3) Sol. Steiner –N; (4) Sol. Steiner –50 % N, –Fe; (5) Sol. Steiner –50 % N, –50 % Fe; (6) Sol. Steiner	
---	--

-50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner, Testigo. Muestreo 2 (56 ddt).	86
Cuadro 3.5. Estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la prueba, y nueve clases de salida: (1) Solución Steiner -N, -Fe; (2) Sol. Steiner -N, -50 % Fe; (3) Sol. Steiner -N; (4) Sol. Steiner -50 % N, -Fe; (5) Sol. Steiner -50 % N, -50 % Fe; (6) Sol. Steiner -50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner, Testigo. Muestreo 3 (67 ddt).	87
Cuadro 3.6. Estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la prueba, y nueve clases de salida: (1) Solución Steiner -N, -Fe; (2) Sol. Steiner -N, -50 % Fe; (3) Sol. Steiner -N; (4) Sol. Steiner -50 % N, -Fe; (5) Sol. Steiner -50 % N, -50 % Fe; (6) Sol. Steiner -50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner, Testigo. Muestreo 4 (78 ddt).	88
Cuadro 3.7. Entrenamiento y prueba promedio de 10 particiones para crear el clasificador con redes neuronales artificiales y tres escenarios de entrada e identificar nueve clases de salida o niveles de fertilización con base en la textura y color de hojas de frijol: (1) Solución Steiner -N, -Fe; (2) Sol. Steiner -N, -50 % Fe; (3) Sol. Steiner -N; (4) Sol. Steiner -50 % N, -Fe; (5) Sol. Steiner -50 % N, -50 % Fe; (6) Sol. Steiner -50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner, Testigo. Muestreo 1 (35 ddt), Muestreo 2 (56 ddt), Muestreo 3 (67 ddt) y Muestreo 4 (78 ddt).	89
Cuadro 3.8. Estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH y DDT en la prueba, y nueve clases de salida: (1) Solución Steiner -N, -Fe; (2) Sol. Steiner -N, -50 % Fe; (3) Sol. Steiner -N; (4) Sol. Steiner -50 % N, -Fe; (5) Sol. Steiner -50 % N, -50 % Fe; (6) Sol. Steiner -50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner,	

Testigo. Datos en conjunto de prueba y entrenamiento del Muestreo 1 (35 ddt),
Muestreo 2 (56 ddt), Muestreo 3 (67 ddt) y Muestreo 4 (78 ddt). 90

ANÁLISIS TEXTURAL Y DE COLOR DE IMÁGENES DIGITALES DEL CULTIVO DE FRIJOL (*Phaseolus vulgaris* L.) EN CONDICIONES DE DEFICIENCIA DE HIERRO Y MANGANESO USANDO CLASIFICADORES CON REDES NEURONALES ARTIFICIALES



INTRODUCCIÓN GENERAL

El desarrollo de herramientas alternativas a métodos tradicionales de diagnóstico nutrimental de un cultivo se ha llevado a cabo al partir de la premisa de la utilidad que tiene la aplicación de técnicas computacionales en el ámbito agrícola sin que se limiten solo a este aspecto. Se tiene en cuenta que la agricultura es un negocio productivo único, que depende tanto de factores climáticos como económicos, dentro de los cuáles se pueden mencionar: suelo, clima, cultivo, riego, fertilización, temperatura, lluvia, cosecha, plagas, entre otros (Majumdar *et al.* 2017); se entiende entonces que tales factores sean estudiados en tantos enfoques como sea posible para encontrar el parámetro óptimo que se adecue a las necesidades recomendables de un cultivo en particular para que éste exprese su rendimiento potencial óptimo.

El objetivo de cada estudio es entonces la generación de información que ayude en la toma de decisiones, de la que dependerá esencialmente alcanzar el objetivo, que es la producción. En términos básicos la toma de decisiones es el proceso de definición de problemas, recopilación de datos, generación de alternativas y selección de un curso de acción y se define como “el proceso para identificar y solucionar un curso de acción para resolver un problema específico” (Stoner *et al.* 2003). Debe destacarse que en el ámbito de la agricultura, una característica de este proceso es que sea llevado a cabo con prontitud.

A lo largo del tiempo se han desarrollado diversas herramientas que se engloban dentro de conceptos como el de agricultura de precisión, y que precisamente acortan el tiempo en el proceso de la toma de decisiones. De manera particular, en el factor de la nutrición un ejemplo de esto es el uso del índice de verdor en herramientas como el SPAD (Soil Plant Analysis Development) para llevar a cabo un manejo de la fertilización nitrogenada en diversas especies como arroz, maíz, jitomate y algodón, entre otras (Xiong *et al.* 2015), o bien en combinación con métodos que emplean el uso del color de las hojas en el caso del cultivo de arroz (Maiti *et al.* 2004). En ambos casos se demostró que existe una correlación entre la información generada por estos métodos y la fertilización nitrogenada del cultivo. Evidentemente la relación del nitrógeno con el contenido de clorofila es clave en los resultados obtenidos como se ha demostrado en investigaciones como la de Datey y Kate (2017).

Sin embargo, en investigaciones como la llevada a cabo por Sartin *et al.* (2014) también se han identificado deficiencias de otros macronutrientes además del nitrógeno (N), como fósforo (P) y en cultivos como el algodón en una etapa temprana. En este caso se tomó en cuenta la variación de color en imágenes provenientes de las hojas del algodón y se procesó esta información usando redes neuronales artificiales. Se concluyó la eficiencia de las redes neuronales en la identificación de las deficiencias de N y P. Años atrás, Hetzroni *et al.* (1994) emplearon clasificadores para identificar deficiencias de hierro (Fe), zinc (Zn) y N en hojas de lechuga, también con el uso de redes neuronales artificiales o más recientemente, en trabajos relacionados como el de Ji-Yong *et al.* (2011).

Por supuesto, la aplicación de la estadística en el análisis de la información fue decisiva en los resultados de los ejemplos mencionados, así que, mientras el uso de la estadística paramétrica ha sido aplicado de manera exitosa, la complejidad de aquella información que no cumple con los supuestos propios de esta rama encontró solución en el uso de la estadística no paramétrica. Son precisamente las redes neuronales artificiales el ejemplo del otro enfoque del análisis de datos como se ha descrito en los párrafos anteriores. Se ha demostrado además, su eficacia en otros campos de la investigación. Según Paliwal y Kumar (2009), desde las últimas dos décadas, las redes neuronales se habían estado usando como alternativa teórica a los modelos estadísticos paramétricos. Y, aunque las redes neuronales tuvieron su origen en la neurobiología matemática, los modelos prácticos bastante simplificados en la actualidad han sido movidos hacia el campo de la estadística. Pitarque *et al.* (2000) señalan que aunque los resultados no son concluyentes; en general, las redes neuronales han mostrado una capacidad clasificatoria igual o superior que otras técnicas estadísticas, con la ventaja de poderse utilizar independientemente del cumplimiento de los supuestos teóricos relativos a tales métodos.

De acuerdo con los autores arriba mencionados, las redes neuronales artificiales han sido aplicadas a problemas comúnmente abordados en la estadística clásica como predicción y clasificación (a través de las llamadas redes hetero-asociativas: perceptrón multi-capas y redes de función base radial), reducción de la dimensionalidad (a través de las llamadas redes auto-asociativas: Hopfield, Kohonen, etc.), series temporales, entre otros, sobre todo cuando se cuenta con datos complejos con un gran número de variables y de dónde se desea obte-

ner información, encontrar patrones y definir tendencias para explicar un fenómeno. Además, Paliwal y Kumar (2009) concluyen que dentro de las ventajas de mayor importancia, es que las redes neuronales pueden aproximarse de manera automática a funciones matemáticas no lineales, y que esta característica es útil cuando la relación entre las variables no es conocida o bien, es complicada y es difícil de modelar.

En la presente investigación se hace énfasis en un problema de deficiencias nutrimentales de Fe o manganeso (Mn) en el cultivo de frijol. Respecto a la deficiencia de Fe, se sabe que es un problema nutrimental frecuente debido a la lixiviación de los fertilizantes del suelo, aireación pobre del suelo, bajo contenido de materia orgánica, temperatura, pH elevado o contenido de bicarbonatos (HCO_3^-) en el agua e interacciones con altos niveles de Mn y calcio (Ca). La mayoría de las plantas deficientes del Fe son amarillas y enanas, con síntomas en las hojas cercanas a la superficie de la planta debido a la inmovilidad del Fe y a la translocación deficiente, lo que resulta en una clorosis intervenal típica de esta deficiencia (Ozores-Hampton, 2013). De acuerdo con Prasad (2003), las deficiencias de Fe generalmente ocurren no por un bajo contenido de Fe en el suelo, sino por varios factores climáticos, edáficos y de la planta que afectan la disponibilidad, absorción o uso metabólico de este micronutriente. La deficiencia de Fe en los cultivos es uno de los desórdenes nutrimentales que se presentan en suelos calcáreos y arenosos de las regiones áridas y semiáridas del mundo. Se ha estimado que cerca de un tercio de los suelos en el mundo son calcáreos y tienen un alto potencial para las deficiencias de Fe. Esto se ha convertido en un problema mundial, causando una pérdida del ren-

dimiento en numerosos cultivos. Este mismo autor agrega que los factores climáticos tienen una influencia muy grande en la ocurrencia de la deficiencia de Fe en los cultivos en condiciones de campo y los más importantes son la temperatura y el contenido de humedad. Al respecto, de los suelos calcáreos, Marschner y Römheld (1994), en plantas que utilizan la estrategia I de absorción de Fe, como es el caso del frijol, dichos suelos constituyen una limitante, por lo que la planta tiene que liberar reductantes/quelatores, particularmente excreción de protones (H^+), acción que es incrementada bajo condiciones de deficiencia de Fe. Debido a esto, hay una acidificación de la rizósfera, con lo que se intenta aumentar la capacidad reductora del Fe para convertirlo a soluble.

Dado que la absorción de Fe y la translocación de la raíz a los brotes es un proceso activo, la temperatura influencia la ocurrencia de la deficiencia de Fe (Fageria *et al.* 1989). Del mismo modo puede influenciar la severidad con la que esta deficiencia se presente en las siguientes formas (Prasad, 2003):

- a. Las bajas temperaturas del suelo reducen el crecimiento radical y la actividad metabólica, por lo tanto, se reduce la absorción de Fe.
- b. Las bajas temperaturas podrían reducir la producción de fitosideróforos, lo que afectaría la movilización y absorción del Fe del suelo.
- c. Las bajas temperaturas podrían incrementar los niveles de HCO_3^- en el suelo y la severidad de la clorosis se vería incrementada al aumentar la solubilidad de CO_2 en los suelos.

- d. Las temperaturas altas en el suelo provocarían un decremento en la absorción de Fe al incrementarse la descomposición microbiana de los fitosideróforos.
- e. Las temperaturas altas podrían incrementar los niveles de HCO_3^- así como la severidad de la clorosis al estimular la actividad microbiana y la producción de CO_2 .

Tang *et al.* (1991) ya habían también concluido que las concentraciones altas de HCO_3^- inducen los síntomas visuales de la clorosis férrica, además de causar un decremento en la fotosíntesis neta de dicotiledóneas como el frijol. Las leguminosas, agregan, son particularmente afectadas por una supresión en la nodulación y en la fijación de N_2 . Jolley *et al.* (1992) también refuerzan la afirmación acerca de los HCO_3^- y agregan que tienen un efecto negativo en plantas que utilizan la estrategia I de absorción de Fe. Además, indican, perjudican la reducción ligada a la membrana al expulsar H^+ y evitan la acidificación del espacio intercelular entre la pared celular y la membrana. Al ocurrir esto, la reducción del Fe^{3+} decrece y por lo tanto, la absorción del Fe decrece. Otro factor que también podría inducir a la deficiencia de Fe es la alta concentración de nitratos en el suelo, sin embargo, no existe evidencia suficiente para explicar este fenómeno (Vasconcelos y Grusak, 2013).

Respecto a la humedad del suelo, Prasad (2003) y Voss (1998) afirman que cuando los niveles de esta son altos, hay un efecto fuerte en la clorosis férrica que tiene efecto en el metabolismo de la planta. Varios estudios indican que el exceso de riego o periodos prolongados de humedad en suelos calcáreos resultan en clorosis férrica debido a un incremento en los HCO_3^- . Los

incrementos en la clorosis siguen a periodos de riego debido a los altos niveles de HCO_3^- en el agua de riego. Además, la presencia de estos aniones eleva el pH del suelo y los bajos contenidos de Fe en suelos con aireación deficiente causada por el exceso de agua destruye muchas de las raíces pequeñas y reduce la capacidad de absorción de todo el sistema radical, lo que podría llevar a la clorosis férrica. Sin embargo, en suelos anegados, las especies Fe^{3+} , particularmente amorfas, son reducidas a Fe^{2+} por las bacterias anaeróbicas que usan Fe^{3+} como aceptoras de electrones durante la respiración ($Fe(OH)_3 + 3H^+ + e^- \rightarrow Fe^{2+} + 3H_2O$). El resultado es un alto contenido de Fe soluble (Fe^{2+}) en suelos anegados, lo que conlleva a niveles tóxicos de Fe ($450 \text{ mg}\cdot\text{kg}^{-1}$). Sin embargo, bajo condiciones normales, estos organismos constituyen una ventaja ya que mejoran la carga apoplásmica con Fe inorgánico. Dependiendo del estado nutricional del cultivo, esta carga de Fe es utilizada en diferente grado de acuerdo con el estatus nutrimental de Fe. En condiciones de deficiencia, mucho del Fe apoplásmico se moviliza y trasloca a los brotes a diferencia de plantas con Fe suficiente (Marschner y Römheld, 1994). Voss (1998) además agrega que, en suelos alcalinos o calcáreos, la actividad microbiana podría producir suficiente CO_2 que reaccione con el agua para formar HCO_3^- , los cuales, al ser absorbidos por la planta, inmovilizan el Fe dentro de la planta, resultando en la deficiencia de dicho elemento. Otro factor que interviene en la incidencia de clorosis en el cultivo, es el pH. A medida que el pH incrementa (7.4 a 8.5), la concentración de Fe^{2+} en la solución decrece. Por otro lado, al disminuir el pH (debajo de 5), podría presentarse toxicidad por Fe. En esta condición podría además presentarse una deficiencia de Mn. Respecto a la materia orgánica, ésta mejora la disponibilidad de Fe al combinarse con éste. De hecho, la misma materia orgánica

nica se convierte en una fuente de este elemento al descomponerse (Voss, 1998). En cuanto a la relación con otros elementos, niveles altos de P, Zn, Mn y cobre (Cu) podrían inducir también una deficiencia de Fe en los cultivos (Fageria *et al.* 1990).

Por otro lado, la deficiencia de Mn se puede identificar por una clorosis intervenal, mientras que las nervaduras permanecen verdes. Al igual que la deficiencia de Fe, ésta se presenta en las hojas jóvenes, pero en la deficiencia de Mn, la diferencia entre coloración de nervaduras y espacio intervenal es difusa (Hong *et al.* 2010). No obstante, Voss (1998) menciona que ambas deficiencias podrían confundirse en cultivos como la soya.

En suelos anegados, se presenta la toxicidad por Mn. Condiciones de humedad excesivamente prolongadas, bajo las cuales haya lixiviación, derivaría en deficiencia de Mn. Por otro lado, en suelos secos, la especie de Mn prevalente es insoluble (Mn^{4+}), por lo que hay condiciones para que ocurra una deficiencia de Mn. Esta reducción tiene lugar cuando el potencial redox es menor a 300 mV. (Chatzistathis, 2014; Voss, 1998). Cuando se da el exceso de humedad y la alta actividad microbiana, hay una aireación del suelo deficiente y por lo tanto, la concentración de oxígeno (O_2) disminuye. Una concentración de materia orgánica alta y temperaturas altas favorecen la alta actividad microbiana, lo que reduce el suministro de O_2 a la planta. Como resultado, el óxido de manganeso (MnO) se reduce a Mn soluble (Mn^{2+}), incrementándose la disponibilidad de este elemento. Bacterias del género *Pseudomonas* y *Citrobacter* son

capaces de oxidar el Mn soluble (a través de procesos enzimáticos y no enzimáticos). Otro género, el *Arthrobacter* puede tener el mismo efecto en la oxidación del Mn en el rango de pH de 5.7 a 7.5 (Chatzistathis, 2014; Marschner, 2012).

Respecto al pH, la disponibilidad de Mn en el cultivo decrece a medida que el pH aumenta, por lo que una deficiencia podría ocurrir a un valor de pH superior a 6.5. Pero debajo de valores de pH inferiores a pH a 5.5, la cantidad de Mn soluble se multiplica 100 veces por cada unidad de pH que disminuya, ocurriendo una toxicidad por Mn y una deficiencia de Fe. En cuanto a los niveles de otros elementos, un alto contenido de P, Ca, Zn y Cu podrían inducir una deficiencia de Mn (Voss, 1998).

Los contenidos altos de iones fosfóricos en los suelos pueden llevar a la deficiencia de Mn debido a la formación de una forma insoluble de Mn. Por otro lado, la textura del suelo también influye en la absorción de este micronutriente; en particular, el Mn en suelos arenosos puede lixiviarse con facilidad mientras que esto no ocurre en suelos arcillosos (Chatzistathis, 2014; Alifragis y Papamichos, 1994).

En condiciones de campo, la deficiencia de Mn ocurre como un desorden latente con síntomas visuales no aparentes. Por esta razón, la magnitud de la deficiencia de Mn, constituye un problema mayor del que se esperaría cuando los síntomas visuales aparecen. El fotosistema II (PSII) es el más afectado con la deficiencia de Mn. De aquí que la medición del nivel de fluorescencia de la clorofila sea un indicador de referencia para ubicar la magnitud de la defi-

ciencia de este micronutriente. El PSII es un complejo proteico de pigmentos y un componente clave en las reacciones de luz en el proceso de la fotosíntesis (Shi *et al.* 2012). El Mn es un constituyente del complejo de liberación de oxígeno del PSII y de la isoenzima peróxido dismutasa con Mn (MnSOD), mientras que también participa en la fotólisis del agua y juega un papel importante en la fotosíntesis en el transporte de electrones en el PSII. Además, este elemento también está involucrado en el metabolismo de carbohidratos, en el ciclo de Krebs, así como en la ruta del ácido shikímico que está relacionado con la producción de fenoles, que a su vez incrementan la resistencia de las plantas a hongos y ataques de insectos.

En cuanto a la selección del cultivo para propósitos de la presente investigación, se sabe que el frijol se cultiva, por lo general, en zonas con condiciones que incrementan la susceptibilidad a la clorosis férrica (Lara Mireles, 2004), o bien, toxicidades por Mn dada la relación entre estos dos micronutrientes como se ha demostrado previamente en trabajos como el de Kohno y Foy (1983) por citar algún ejemplo. Para relacionar la problemática y presentar un contexto, en nuestro país, Zacatecas es el estado en el que se produce el 29.9 % del frijol y que representa, además, el 34.8 % de la superficie sembrada a nivel nacional con este cultivo (SIAP, 2016). En 2015, las entidades federativas de Zacatecas (29.9 %), Durango (11.5 %), Sinaloa (8.7 %), Chihuahua (9.9 %), Chiapas (6.2 %), San Luis Potosí (4.6 %) y Guanajuato (5.7 %) concentraron el 76.6 % la cosecha nacional de frijol (FIRA, 2016).

De acuerdo con el Censo Agropecuario 2007 (INEGI, 2013), los suelos en Zacatecas son catalogados principalmente como leptosoles, calcisoles y phaeozems (29.5, 18.3 y 19.6 %, respectivamente), además, en este mismo censo se registra que el frijol supera en volumen de producción al cultivo principal en México: el maíz. Cabe considerar que la importancia del frijol no es únicamente agrícola sino económica, social y cultural. Se encuentra también, que este cultivo se realiza en 32 estados de la República, en zonas de riego, temporal, y bajo condiciones de humedad residual; sin embargo, la producción de frijol es muy vulnerable a las condiciones climatológicas que prevalecen durante el ciclo productivo, debido a que aproximadamente 87.0 % de la superficie nacional destinada a este cultivo se ubica en áreas de temporal. Los problemas nutricionales del cultivo obedecen al monocultivo, lo que causa agotamiento de la tierra por pérdida de fertilidad, erosión e incidencia de plagas y enfermedades. A su vez, esto se traduce en mayor inversión de fertilizantes y plaguicidas. Otro factor a considerar, se reafirma, es que no toda la superficie destinada para el cultivo de frijol es apta para sembrarlo, ya que carece de potencial agroecológico.

OBJETIVOS

Objetivo General

- Identificar deficiencias de hierro y manganeso en el cultivo de frijol con el uso de análisis digital de imágenes, redes neuronales artificiales, y árboles de predicción.

Objetivos Específicos

- Usar dos modelos de predicción del contenido de Fe en el cultivo de frijol, redes neuronales Bayesianas regularizadas (BRNN) y árboles de clasificación utilizando como herramienta lecturas SPAD a través del tiempo durante la etapa vegetativa de cultivo.
- Crear clasificadores con base en las redes neuronales artificiales (RNA) y características texturales y de color para identificar deficiencias de Fe y Mn en hojas de frijol en etapa reproductiva de prefloración (R5) y en la etapa reproductiva de floración (R6).
- Proponer clasificadores de deficiencias de Fe en el cultivo de frijol y el efecto de tres niveles de N con el uso de RNA y características texturales y de color en cuatro etapas de desarrollo del cultivo.

HIPÓTESIS

- Las características texturales y de color forman patrones específicos en las hojas del cultivo de frijol que varían de acuerdo con el nivel de nutrición del mismo. Estos patrones pueden ser identificados mediante técnicas computacionales como las redes neuronales artificiales.
- La correlación entre el contenido de clorofila y la concentración de Fe soluble permite estimar el estatus nutrimental de este micronutriente en el cultivo de frijol mediante la obtención del índice de verdor calculado por el SPAD 502 (Minolta Corp.).
- Dado el papel principal del N en la conformación de la molécula de clorofila, el estatus nutrimental del N afectaría la identificación de las deficiencias de Fe con el uso de clasificadores creados con redes neuronales artificiales.

LITERATURA CITADA

- Alifragis, D. and N. Papamichos. 1994. Fertility and Fertilization of Forest Soils. Aristotle University, Thessaloniki, Greece, 235 pp.
- Chatzistathis, T. 2014. Plant Responses to Iron, Manganese, and Zinc Deficiency Stress *in* Emerging Technologies and Management of Crop Stress Tolerance Volume 1: Biological Techniques. Pages 293–311.
- Datey, R. and D. Kate. 2017. A Review On: Detection of Chlorophyll and Nitrogen in Plants using Digital Image Processing. *International Journal of Advanced Engineering, Management and Science* 3: 108 – 110.
- Fideicomisos Instituidos en Relación con la Agricultura (FIRA). 2016. Panorama agrolimentario. Frijol 2016. Dirección de Investigación y Evaluación Económica y Sectorial. México.
- Hetzroni, A., G. E. Miles, B. A. Engel, P. A. Hammer, and R. X. Latin. 1994. Machine vision monitoring of plant health. *Advances in Space Research* 14(11):203-212.
- Hong, E., Q. Ketterings, and M. McBride. 2010. Manganese. Fact Sheet 49. Agronomy Fact Sheet Series. Nutrient Management Spear Program. Cornell University Cooperative Extension.
- Instituto Nacional de Estadística y Geografía (INEGI). 2013. Censo Agropecuario 2007. El cultivo del frijol en Zacatecas. México.
- Ji-yong, S., Z. Xiao-bo, Z. Jie-Wen, M. Han-Ping, W. Kai-Liang, C. Zheng-Weil, and H. Xiao-Weil. 2011. Diagnostics of nitrogen deficiency in mini-cucumber plant by near infrared reflectance spectroscopy. *African Journal of Biotechnology* 10(85):19687-19692.
- Jolley, V. D, D. J. Fairbanks, W. B. Stevens, R. E. Terry, and J. H. Off. 1992. Root iron-reduction capacity for genotypic evaluation of iron efficiency in soybean. *Journal of Plant Nutrition* 15: 1679-1690.

- Kohno, Y. and C. D. Foy. 1983. Manganese toxicity in bush bean as affected by concentration of manganese and iron in the nutrient solution. *Journal of Plant Nutrition* 6(5):363-386.
- Lara-Mireles, J. L., R. Vázquez-Alvarado, E. Olivares-Sáenz y J. F. Pissani-Zúñiga. 2004. Tolerancia a clorosis férrica de diferentes cultivares de frijol en suelos calcáreos. *Revista Fitotecnia Mexicana* 27(1):43-47.
- Maiti, D., D.K. Das, T. Karak, and M. Banerjee. 2004. Management of nitrogen through the use of leaf color chart (lcc) and soil plant analysis development (SPAD) or chlorophyll meter in rice under irrigated ecosystem. *The Scientific World Journal* 4: 838–846.
- Majumdar, J., S. Naraseyappa, and S. Ankalaki. 2017. Analysis of agriculture data using data mining techniques: application of big data. *Journal of Big Data* 4:20.
- Marschner, P. 2012. *Marschner's Mineral Nutrition of Higher Plants*. Academic Press. London.
- Marschner, H. and V. Römheld. 1994. Strategies of plants for acquisition of iron. *Plant Soil* 165: 261-274.
- Ozores-Hampton, M. 2013. Effective strategies to correct iron deficiency in Florida vegetable crops. *Hortechonology* 23(5): 548-552.
- Paliwal, M. and U. A. Kumar. 2009. Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications* 36: 2–17.
- Pitarque, A., J. C. Ruiz, and J. F. Roy. 2000. Las redes neuronales como herramientas estadísticas no paramétricas de clasificación. *Psicothema* 12(2): 459-463.
- Prasad, P. V. V. 2003. Iron Chlorosis. *Encyclopedia of applied plant sciences*. pp. 649-656.
- Sartin, M. A., A. C. R. Da Silva, and C. Kappes. 2014. Image segmentation with artificial neural network for nutrient deficiency in cotton crop. *Journal of Computational Science* 10 (6): 1084-1093.

- Servicio de Información Agroalimentaria y Pesquera (SIAP). 2016. Atlas Agroalimentario 2016. México.
- Shi, L., M. Hall, C. Funk, and W. P. Schröder. 2012. Review: Photosystem II, a growing complex: Updates on newly discovered components and low molecular mass proteins. *Biochimica et Biophysica Acta* 1817: 13–25.
- Stoner, J. A. F., R. E. Freeman y D. R. Gilbert Jr. 2003. Administración. Ed. 6ta. Prentice-Hall. México.
- Tang, C., A. D. Robson, M. J. Dilworth, and J. Kuo. 1992. Microscopic evidence on how iron deficiency limits nodule initiation in *Lupinus angustifolius* L. *New Phytologist* 121: 457-467.
- Vasconcelos, M. W. and M. A. Grusak. 2013. Morpho-physiological parameters affecting iron deficiency chlorosis in soybean (*Glycine max* L.). Publications from USDA-ARS / UNL Faculty. Paper 1302.
- Voss, R. 1998. Micronutrients. Department of Agronomy. Iowa State University. (http://www.agronext.iastate.edu/soilfertility/info/Micronutrients_VossArticle.pdf)
- Xiong, D., J. Chen, T. Yu, W. Gao, X. Ling, Y. Li, S. Peng, and J. Huang. 2015. SPAD-based leaf nitrogen estimation is impacted by environmental factors and crop leaf characteristics. *Scientific Reports* 5:13389.

C A P Í T U L O I

MODELOS DE PREDICCIÓN Y CLASIFICACIÓN DE DEFICIENCIAS DE HIERRO EN EL CULTIVO DE FRIJOL (*Phaseolus vulgaris* L.) CON REDES NEURONALES BAYESIANAS REGULARIZADAS Y ÁRBOLES DE CLASIFICACIÓN

Edgar García-Cruz¹, Manuel Sandoval-Villa¹, José Alfredo Carrillo-Salazar¹, Paulino Pérez-Rodríguez¹, Jorge Dionisio Etchevers-Barra¹, Antonia Macedo-Cruz¹

¹Colegio de Postgraduados, Montecillo Estado de México

1.1 RESUMEN

Los métodos de detección de deficiencias de hierro (Fe) en cultivos como el frijol constituyen una herramienta valiosa en la toma de decisiones ya que pueden utilizarse para predecir el estado nutricional de las plantas en etapas tempranas. En esta investigación se usaron redes neuronales Bayesianas regularizadas (BRNN) y árboles de clasificación para llevar a cabo la predicción de dichas deficiencias basados en lecturas del SPAD 502, el cual se empleó para medir el índice de verdor de las hojas en el frijol. Se llevó a cabo un experimento con ocho tratamientos con diferentes variaciones en la concentración de Fe (0, 20, 40, 60, 80, 100, 150 y 200 %) en la solución nutritiva. Durante 7 semanas se tomaron las mediciones promedio del índice de verdor de los tres folíolos de cinco repeticiones correspondientes a los ocho tratamientos y posteriormente los datos fueron utilizados para ajustar los modelos estadísticos antes mencionados. Con las BRNN, la correlación entre valores observados y predichos fue de 0.77 para el conjunto de datos en entrenamiento y de 0.54-0.71 para prueba. Para el caso de los árboles de clasificación, en el conjunto de entrenamiento el porcentaje de clasificaciones correctas fue 56.25 % y disminuyó casi 30 % cuando se llevó a cabo el procedimiento de validación. Por lo que para la presente investigación, el uso de BRNN constituye una herramienta valiosa para la predicción de deficiencias tempranas de Fe en el cultivo de frijol.

Palabras clave: SPAD, clorosis férrica, detección, nutrición, índice de verdor.

**CLASSIFICATION AND PREDICTION MODELS OF IRON DEFICIENCIES IN BEAN
LEAVES (*Phaseolus vulgaris* L.) USING BAYESIAN REGULARIZED NEURAL NET-
WORKS AND CLASSIFICATION TREES**

Edgar García-Cruz¹, Manuel Sandoval-Villa¹, José Alfredo Carrillo-Salazar¹, Paulino Pérez-
Rodríguez¹, Jorge Dionisio Etchevers-Barra¹, Antonia Macedo-Cruz¹

¹Colegio de Postgraduados, Montecillo Estado de México

1.2 SUMMARY

The detection methods for iron (Fe) deficiencies in crops such as beans are valuable tools in decision making because those methods can be used to predict the nutritional status of plants in early stages of the crop cycle. For this research, Bayesian regularized neural networks (BRNN) and classification trees were used to predict the deficiencies above mentioned based on a SPAD 502 readings that were used to measure the greenness index in bean leaves. An experiment was set up using eight treatments with variations in Fe concentration in the nutrient solution (0, 20, 40, 60, 80, 100, 150, and 200 %). For seven weeks, the average green index measurements of the three leaflets of five replicates corresponding to the eight treatments were taken and subsequently the data were used to adjust the statistical models mentioned above. With BRNN, the correlation between observed and predicted values was 0.77 for the data set in training set and 0.54-0.71 for the validation set. In the case of classification trees, in the training set the percentage of correct classifications was 56.25 % and decreased almost 30 % when the validation procedure was carried out. As for the present research, the use of BRNN constitutes a valuable tool for the prediction of early deficiencies of Fe in bean cultivation.

Keywords: SPAD, iron chlorosis, detection, nutrition, green index.

1.3 INTRODUCCIÓN

La deficiencia de hierro (Fe) es un problema mundial en la producción de cultivos en suelos calcáreos. Dentro de los cultivos afectados por la deficiencia de Fe, el frijol (*Phaseolus vulgaris*) es uno de ellos ya que se encuentra dentro de la categoría de susceptible a la clorosis férrica (Hansen *et al.* 2007 y Zaiter *et al.* 1992). De acuerdo con Clark (1991), este problema puede reducir hasta el 100 % del rendimiento del cultivo. La clorosis férrica se ve expresada en la planta una vez que el estado de la deficiencia es avanzado, y esta decoloración de las hojas se debe a un desequilibrio entre los contenidos de clorofila y los carotenos por la reducción del número de tilacoides y granas; de esta forma, hay una disminución de los pigmentos que recogen la luz: clorofilas *a* y *b* y carotenos (Soldatini *et al.* 2000). A nivel molecular, el Fotosistema I (PSI) es el primer afectado por la deficiencia de Fe, por lo que disminuye el número de estos complejos, lo que resulta en un cuello de botella en el flujo de electrones necesarios en la fotosíntesis (Yadavalli *et al.* 2012; Strzepek y Harrison, 2004). Una vez que el daño a nivel molecular se ha dado, ha ocurrido ya una disminución al rendimiento del cultivo, por lo que las acciones correctivas que pudieran llevarse a cabo solo tendrían un efecto amortiguador. De aquí la importancia de llevar a cabo la detección a tiempo del problema para ejecutar acciones que puedan evitar un decremento en el rendimiento más allá de un umbral económico.

Uno de los métodos para realizar el diagnóstico nutrimental en el cultivo es el uso de los análisis químicos en laboratorio, sin embargo, a pesar de la precisión que éstos puedan tener, su utilización constituye una inversión de tiempo que puede evitarse con métodos de detección

in situ. En el caso de elementos como el nitrógeno (N) existen aparatos específicos para la estimación del contenido de este elemento en el cultivo, no así para el caso de Fe. Un instrumento que permite estimar el contenido de N es el medidor de contenido de clorofila SPAD 502 (Minolta Corp.). Markwell *et al.* (1995) sugieren que probablemente la determinación de clorofila en la hoja por unidad de área es relativamente independiente de las especies, por lo que la relación de la clorofila con los valores SPAD sería la misma. El análisis de Fe²⁺ en la hoja se ha llegado a tomar como un indicador del nivel de abastecimiento de hierro en la planta (Razeto y Valdés, 2006) e investigadores como Köseoğlua y Açikgöza (1995) han llevado a cabo experimentos donde se ha encontrado una alta correlación entre la concentración de clorofila y la de Fe²⁺ en la hoja de duraznero (*Prunus persica* L.), tanto en tejido fresco como en tejido seco. Lombard *et al.* (2010) han llevado a cabo trabajos usando el SPAD 502 para la estimación de Fe en especies forestales como el álamo y en árboles frutales como perales se usó este equipo para estimar el contenido de clorofila (López-Millán *et al.* 2001). Se tomaron muestras de hojas con diferentes grados de deficiencias de Fe y dónde se habían tomado previamente lecturas SPAD (Lombard *et al.* 2010). Posteriormente, la clorofila se extrajo con acetona al 100 % y en presencia de ascorbato de sodio dado que se supone, existe correlación entre el contenido de Fe y de clorofila. De acuerdo con Lombard *et al.* (2010), el SPAD 502, comparado con otros métodos de extracción caros o escalas de evaluación visual, este medidor de clorofila puede ser considerado como un indicador confiable de estimación del color, aunque cabe mencionar que no hay estudios hechos específicamente para correlacionar los valores de las lecturas SPAD con estatus nutrimental del Fe en el cultivo de frijol.

Existen diferentes modelos estadísticos que se han utilizado para relacionar las lecturas SPAD con la concentración de clorofila como análisis de regresión (Lombard *et al.* 2010, Hirai *et al.* 2007 y Vesali *et al.* 2015). No obstante, una de las metodologías para el análisis de la información obtenida consiste en el uso de métodos estadísticos no paramétricos como las redes neuronales artificiales tal como se demuestra en el trabajo llevado a cabo por Vesali *et al.*(2015), quienes sistematizaron los datos de valores SPAD, características de color e información obtenida de sensores de imagen para predecir la concentración de N en plantas de maíz utilizando una aplicación móvil y la cual aún no está disponible para uso comercial. El desempeño de los diferentes tipos de redes neuronales es variado; de acuerdo con MacKay (1992), éstas pueden usarse tanto en el contexto de regresión como para clasificación. El mismo autor agrega que esta flexibilidad las hace capaces de descubrir relaciones más complejas en los datos que los modelos estadísticos tradicionales. Por otro lado, la estadística Bayesiana provee un marco unificador para la generación de modelos, evitando el problema de sobreajuste que generalmente presentan este tipo de modelos por el hecho de depender de un gran número de parámetros (Foresee y Hagan, 1997). Los modelos Bayesianos pueden ser utilizados con la finalidad de predecir la clase y clasificar cada caso no visto: un nuevo caso o ejemplo j , caracterizado sólo con los valores \mathbf{x}_j de las variables predictoras, se da una etiqueta clase según la categoría que tiene la máxima probabilidad *a posteriori*. Una propiedad útil de un clasificador Bayesiano es que es óptimo, en el sentido que la tasa esperada de falsos positivos se reduce a un mínimo (Ripley, 1996), lo que constituye una ventaja adicional al problema del sobreajuste al ser un método supervisado. Correa *et al.* (2008) señalan que el objetivo

de la clasificación supervisada es categorizar los casos i dados por ciertas características $\mathbf{x}_i = (x_{i1}, \dots, x_{in})'$ en r clases, c_i , $i=1, \dots, r$, x_{il} denota el valor de la variable x_l observada en el caso i .

Aunque la aplicación de las BRNN en agricultura no es nuevo si lo es el uso de métodos de aprendizaje de máquina en la predicción e inteligencia artificial, para predecir deficiencias de nutrientes. En un documento escrito por Bustos (2005), el autor hace un análisis de la inteligencia artificial en diferentes líneas de trabajo en la agricultura, lo que incluye el tipo de redes arriba mencionado. Dentro de algunos de los campos mencionados se encuentra la ingeniería postcosecha, el control ambiental, maquinaria y mecanización agrícola, ingeniería de riegos y producción vegetal, específicamente en aspectos nutricionales de los cultivos. De hecho, hay aplicaciones del uso de estas redes en investigaciones relacionados con la categorización de estatus nutrimental de algún cultivo. Tal es el caso del trabajo llevado a cabo por Ma *et al.* (2014) en el que se discuten métodos de clasificación del verdor de los cítricos para categorizar en plantas sanas, infectadas por *Huanglongbing* (HLB), deficiencia de Fe y deficiencia de N, utilizando BRNN. Otra forma de categorización de la información es el uso de los árboles de clasificación ya que proporcionan predictores bastante comprensibles en situaciones en las que hay variables que interactúan de forma complicada y de forma no lineal. La variable respuesta Y es categórica, así, puede usarse información teórica para medir que tanto puede aprenderse de esta al saber el valor de otra variable discreta A . Hay dos tipos de predicciones que un árbol de clasificación puede hacer. El primer tipo es predicción puntual de la clase o

categoría. El segundo tipo es la probabilidad de pertenencia a cada una de las clases (Shalizi, 2009).

Para la presente investigación, y de acuerdo a la información anterior, el objetivo es usar dos modelos de predicción del contenido de Fe en el cultivo de frijol, redes neuronales Bayesianas regularizadas (BRNN) y árboles de clasificación utilizando como herramienta lecturas SPAD a través del tiempo durante la etapa vegetativa de cultivo.

1.4 MATERIALES Y MÉTODOS

Establecimiento del experimento, trasplante y aplicación de tratamientos

La siembra de semillas de frijol var. Negro Chapingo, se realizó bajo condiciones de invernadero en el Colegio de Posgraduados, en charolas de poliestireno y con perlita como sustrato durante el mes de junio de 2015. El riego se hizo diario con agua destilada.

Doce días después de la siembra (dds), éstas se trasplantaron bajo condiciones de hidroponía en vasos de poliestireno, continuando el riego con agua destilada. Posteriormente, la preparación de las soluciones nutritivas se llevó a cabo también usando agua destilada y se tomó como base a la solución universal Steiner. Cada uno de los vasos en los que se trasplantó una plántula se tomó como una unidad experimental. A partir de 80 unidades, se dividieron éstas en 8 grupos con diez repeticiones. A cada grupo se le aplicó un tratamiento que correspondió a las variaciones de concentraciones de Fe. Los tratamientos fueron condiciones de deficiencia

(0, 20, 40, 60 y 80 %) y exceso (150 y 200 %), además de la condición óptima o testigo (100 %). Estos se aplicaron siete días después del trasplante (ddt). Durante siete semanas se llevaron a cabo 20 mediciones del índice de verdor por cada tratamiento, se tomaron en cuenta cinco repeticiones por los ocho tratamientos, usando un SPAD 502 (Minolta, Corp.), tales repeticiones corresponden a 5 hojas, una hoja de cada planta, de las cuales se tomaron 5 lecturas por cada uno de los folíolos de la hoja para después obtener un promedio. Posteriormente, se tomó en cuenta el promedio de los tres folíolos. Los valores promedio por tratamiento fueron usados para generar el modelo (160 observaciones). Debe mencionarse que las mediciones se tomaron en la etapa vegetativa.

Predicción de deficiencias de Fe usando redes neuronales Bayesianas regularizadas (BRNN)

Se usó un modelo de BRNN (Figura 1.1). Este modelo fue entrenado a partir de un conjunto de datos con los que se pretende predecir la concentración de Fe en plantas de frijol (y), la que se considera como variable respuesta, que se supone es posible predecir utilizando dos covariables (variables de entrada): valores SPAD (x_1) y ddt (x_2). Se parte del uso de S neuronas de la capa interna, tres en este caso (Figura 1.1), las cuáles procesaron la información recibida de las variables de entrada para posteriormente generar una respuesta por cada una de las neuronas localmente. Los valores de entrada en cada una de las neuronas son combinadas linealmente (u_1, u_2, u_3) y la respuesta de esta operación es transformada de forma no lineal utilizando una función de activación, la cual es una función que mapea las variables de entrada de la recta real al intervalo $(-1, 1)$. En este trabajo se empleó la función de activación tangen-

te hiperbólica, $g(x) = \frac{2}{1+\exp(-2x)}$ (Pérez-Rodríguez *et al.* 2012). Una vez que los datos han sido procesados en cada una de las neuronas ($g(u_1)$, $g(u_2)$, $g(u_3)$), las salidas son combinadas linealmente, asignando un peso específico a cada una de las neuronas para obtener la predicción de la variable respuesta. Los parámetros a estimar incluyen; la media general (μ), los coeficientes de regresión en cada una de las neuronas: $\beta_0^k, \beta_1^k, \beta_2^k, k = 1, \dots, S$ (neuronas), los pesos de cada una de las neuronas: w_1, \dots, w_S , así como la varianza del error σ^2 .

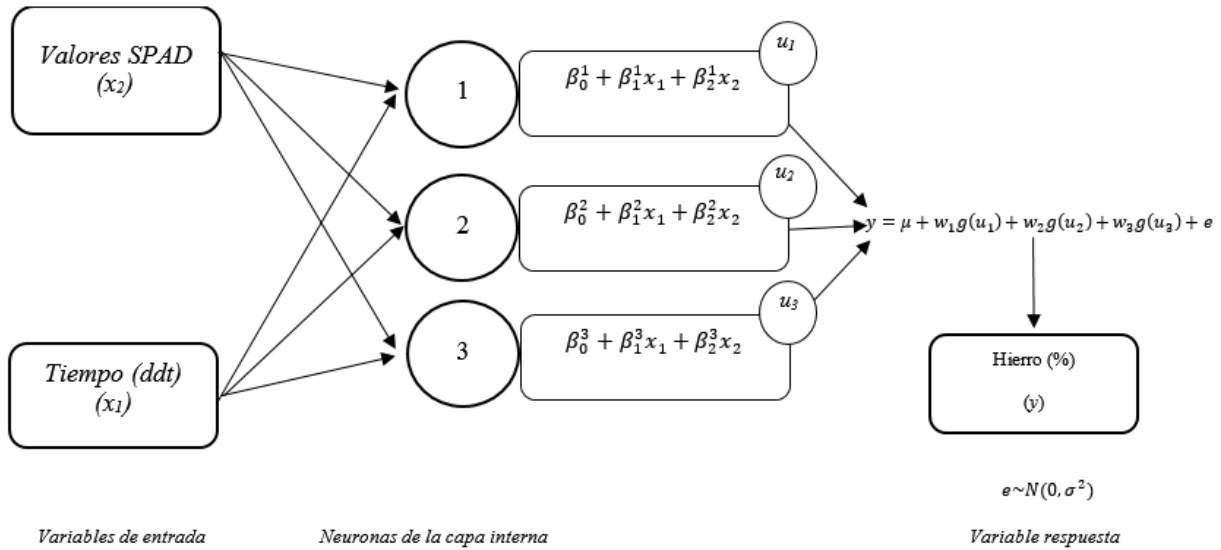


Figura 1.1. Representación gráfica del modelo de red neuronal utilizada para $S = 3$ neuronas.

Para el ajuste del modelo, se utilizó el programa estadístico R (R Core Team, 2015), así como la biblioteca de funciones *brnn* (Pérez-Rodríguez y Gianola, 2016). De acuerdo con estos autores, la biblioteca de funciones *brnn* incluye una función con el mismo nombre que ajusta una red neuronal de dos capas como se describe en MacKay (1992) y Foresee y Hagan (1997), además, usa el algoritmo de Nguyen y Widrow (1990) para asignar pesos iniciales y el algoritmo de Gauss-Newton para llevar a cabo la optimización.

La fase de entrenamiento de la BRNN puede describirse como un proceso de ajuste de los parámetros de forma gradual hasta que el vector de salida predicho aproxime lo más posible con el vector de salida deseado. El algoritmo utilizado para el ajuste del modelo es diferente al de retropropagación de los errores utilizado comúnmente en redes neuronales, en BRNN se minimiza una “suma de cuadrados penalizada”. El algoritmo de entrenamiento se detiene una vez que la “suma de cuadrados penalizada” se estabiliza en tres evaluaciones sucesivas (ver Foresee y Hagan, 1997; Pérez-Rodríguez *et al.* 2012). Una vez que los estimadores de los parámetros de interés han sido obtenidos, es posible evaluar si el número de neuronas utilizados es suficiente para realizar la predicción sin realizar sobreajuste, ya que es posible calcular fácilmente el número efectivo de parámetros, es decir el número de parámetros que desde el punto de vista estadístico son diferentes de cero (Okut *et al.* 2011, Pérez-Rodríguez *et al.* 2012).

Posteriormente se determinó el poder predictivo del modelo utilizando dos esquemas de validación cruzada: 1) conjuntos aleatorios y 2) conjuntos no aleatorios. En el primer caso, de acuerdo a James *et al.* (2013) el procedimiento involucra la división aleatoria de un conjunto de observaciones en un número k de subconjuntos disjuntos (en este caso, ocho) de un tamaño igual aproximadamente. En un segundo procedimiento de validación no aleatorio, se tomaron también ocho grupos de acuerdo a las concentraciones de Fe (0, 20, 40, 60, 80, 100 y 150). En cualquiera de los casos, el conjunto de datos es dividido in k subconjuntos disjuntos $\{S_1, \dots, S_8\}$. Cada uno de estos conjuntos puede utilizarse para medir el poder predictivo de los

modelos. Por ejemplo, utilizando el primer subconjunto, el conjunto de datos se divide en entrenamiento (los datos en S_2, \dots, S_8) y prueba (los datos en S_1). Los modelos se ajustan los datos en el conjunto de entrenamiento y luego se utiliza para predecir la variable respuesta para el conjunto de datos en prueba. Este procedimiento se repite k número de veces (ocho en este caso) y al final del proceso se obtendrán predicciones para el los k grupos de donde se pueden obtener medidas que permiten evaluar el poder predictivo de los modelos, por ejemplo correlaciones de Pearson entre valores observados y predichos o bien tasas de aciertos y errores.

Predicción de deficiencias de Fe usando con árboles de clasificación

De acuerdo con Loh (2011), los árboles de regresión, tanto como de clasificación son métodos de aprendizaje automático para la construcción de modelos a partir de datos. Los árboles de clasificación se construyen a partir de variables dependientes que tienen un número finito de valores sin ordenar y con un error de predicción medido en términos de clasificación errónea.

Por otro lado, para los árboles de predicción, la rutina utilizada fue *tree* en el programa estadístico R (Ripley, 2016). Estos modelos son obtenidos por la partición recursiva del espacio de datos y ajustando a un nivel de predicción simple dentro de cada partición. El mismo autor agrega que, en un problema de clasificación se tiene una muestra de entrenamiento con n observaciones en una variable clase y que toma los valores $1, 2, \dots, k$, y p variables de predicción, x_1, x_2, \dots, x_p . Para la presente investigación, la variable clase es corresponde al porcen-

taje de Fe en la solución nutritiva (y), mientras que son dos las variables de predicción: ddt en que se tomó la medición con el SPAD 502 (x_1), así como las lecturas promedio tomadas por cada hoja (x_2). Para llevar a cabo la predicción se aplica el siguiente algoritmo (Loh, 2011):

- (1) Comenzar en un nodo raíz.
- (2) Para cada variable x ordenada, convertirla a una variable no ordenada x' agrupando sus valores en el nodo en un pequeño número de intervalos. Si x es no ordenada, fijar $x'=x$.
- (3) Llevar a cabo una prueba de independencia *chi* cuadrada de cada variable x' contra y en los datos en el nodo y calcular su significancia.
- (4) Elegir la variable x^* asociada con la x' que tenga la probabilidad de significancia más pequeña.
- (5) Encontrar el conjunto $\{x^* \in S^*\}$ que minimice la suma de los índices Gini (varianza binomial) y la use para dividir el nodo en dos hijos y donde s es un subconjunto de valores tomados por x .
- (6) Si hay convergencia, terminar. De otra forma, aplicar los pasos (2) al (5) para cada nodo hijo.
- (7) Podar el árbol para encontrar el mejor tamaño del mismo.

1.5 RESULTADOS Y DISCUSIÓN

Redes neuronales Bayesianas regularizadas

Se observó que las variables de entrada propuestas pueden predecir el estado nutrimental de Fe en el cultivo de frijol el coeficiente de correlación de Pearson entre valores observados y predichos fue de 0.77 en para el conjunto de datos en entrenamiento. En la Figura 1.2 puede apreciarse la tendencia de la concentración de Fe, la cual es similar al inicio de la toma de las lecturas. Como podría esperarse, los valores SPAD no aumentan mucho a medida que avanzan los días (17 a 25 ddt) en el caso del tratamiento con ausencia total de Fe, por lo que el

índice de verdor es el más bajo respecto al resto de los tratamientos (Figura 1.3). Adams *et al.* (2000) habían ya afirmado que bajo condiciones de deficiencia de Fe, entre otros elementos como el zinc (Zn) y manganeso (Mn), causaban un decremento en la concentración de clorofila debido a que hay una relación directa con la ruta biosintética de ésta siendo elementos esenciales; Morales *et al.* (1991) describen que es en el transporte de electrones en el proceso de la fotosíntesis dónde la ausencia de Fe es causante de la clorosis.

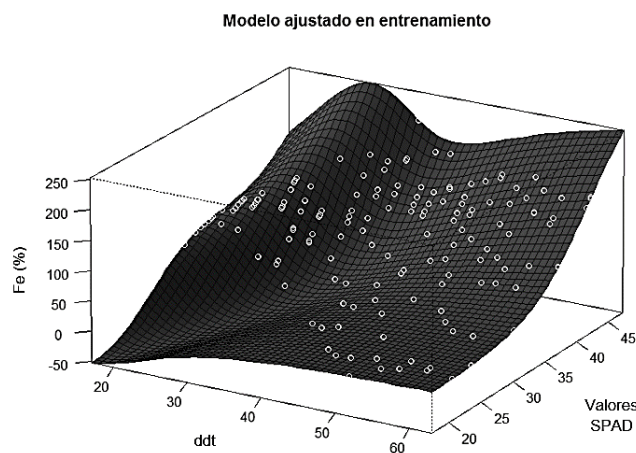


Figura 1.2. Representación gráfica tridimensional que muestra el comportamiento de las diferentes concentraciones de Fe en función del tiempo y el índice de verdor (valores SPAD) en el cultivo de frijol. Estos datos se usaron para el entrenamiento de la BRNN. Los puntos corresponden a las mediciones mientras que en la malla se reflejan los valores que corresponden a las predicciones obtenidas mediante la BRNN con $S = 3$ neuronas.

Por el contrario, si se observa al testigo (Figura 1.2), puede notarse que tiene una tendencia al incremento en el índice de verdor aunque no necesariamente se mantiene ya que hay un decremento de los 30 a 40 ddt. Además, el tratamiento con dosis de Fe más allá del óptimo requerido muestra un comportamiento en el que el índice de verdor es más alto entre los 25 a 30 ddt y posteriormente el decremento es pronunciado, para después registrar los valores SPAD más altos. Este resultado es similar a lo reportado por Kobraee *et al.* (2011) dónde concluye-

ron que los valores SPAD se incrementaron al aumentar la concentración de clorofila y cuando se aplicaron dosis más altas de Fe, entre otros micronutrientes.



Figura 1.3. Apariencia general de la deficiencia de Fe a los 18 ddt en el cultivo de frijol var. Negro Chapingo en dos fases distintas de desarrollo de la deficiencia pero en el mismo tratamiento (-Fe).

Deben también considerarse investigaciones anteriores, como la llevada a cabo por Papasavvas *et al.* (2008) en remolacha, quienes demostraron la correlación entre los valores SPAD y el contenido de clorofila aunque enfocado a la concentración de N (NO_3^-) en la planta. No obstante, debe tomarse en cuenta que tanto el N como el Fe están directamente relacionados con el verdor de las hojas, lo cual también se ha comprobado en trabajos de Ammari *et al.* (2009) en plantas de cítricos.

En validación cruzada, el coeficiente de correlación entre valores observados y predichos fue de 0.54 cuando la partición es no aleatoria. No obstante, cuando se lleva a cabo la validación cruzada con particiones aleatorias, el porcentaje de clasificaciones correctas resultó mayor (0.71). Esto se explica debido a que en los datos de entrenamiento se incluye información de todos los grupos, incluido el que se está validando, no así en el caso anterior de validación cruzada no aleatoria. Cuando se llevó a cabo la validación cruzada con particiones aleatorias,

este procedimiento se repitió 1000 veces, prediciendo aproximadamente 12 % de las observaciones (20). En cada repetición se obtuvo la correlación entre valores observados y predichos y con las mismas se obtuvo una distribución empírica de las mismas. La distribución de las correlaciones se muestra en la Figura 1.4. Usando la distribución empírica de las correlaciones se obtuvo un intervalo de confianza al 95 %, concluyendo que la correlación entre valores observados y predichos está entre 0.68 a 0.75.

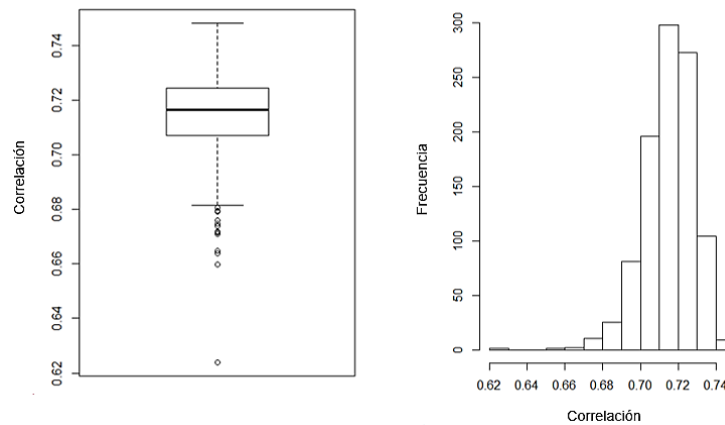
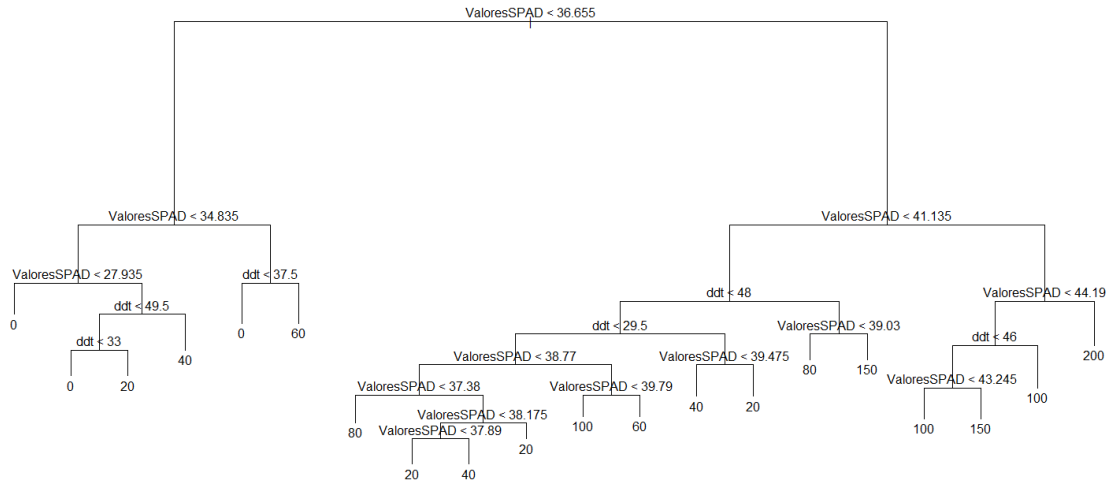


Figura 1.4. Diagrama de cajas y bigotes para las correlaciones obtenidas mediante validación cruzada. Histograma de frecuencias de las diferentes correlaciones entre valores predichos contra valores observados obtenidas después de llevar a cabo el procedimiento de validación cruzada con 8 particiones aleatorias.

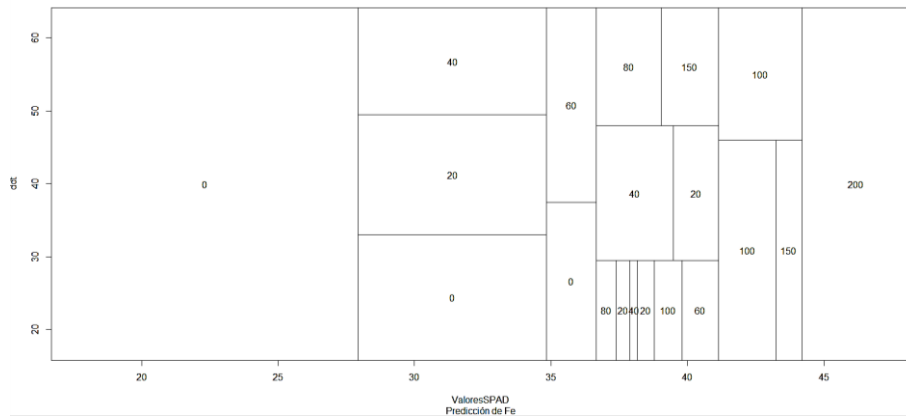
Árboles de predicción y clasificación

El árbol de clasificación obtenido se muestra en la Figura 1.5(a). El nodo principal es asignado a un valor de las lecturas SPAD (36.66), a partir del cual se bifurca en dos nodos hijos, dependiendo si el valor SPAD es mayor o inferior al fijado al nodo principal. Por lo tanto, se observa que las concentraciones de 0 % Fe siempre tendrán valores SPAD inferiores a 34.83, excepto si el día que se tomó la lectura es posterior a los 37 ddt ya que lo concentración mode-

lada será otra. Dicha concentración podrá ser 20, 40 o 60 % si el valor SPAD es superior a 34.83 dependiendo del tiempo en que se haya tomado la lectura.



(a)



(b)

Figura 1. 5. Predicción de Fe en el cultivo de frijol a partir del número de días después del transplante (ddt) y los valores SPAD usando un (a) Diagrama de árbol y (b) Partición binaria del espacio.

Concentraciones como 80, 100, 150 % de Fe siempre tendrían valores SPAD superiores a 36.66, incluso, la concentración de 200 % registraría valores superiores a 44.19. Esta opción representa una alternativa sencilla a para la discriminación entre diferentes clases de deficiencias debido a las condicionantes que presenta para descartar una u otra opción, por lo que

su aplicación en campo sería relativamente sencilla una vez que el método pudiera ser calibrado y mejorado.

Una forma más práctica de visualizar la información es de acuerdo la Figura 4b. Por ejemplo, las concentraciones superiores (200 %) a la óptima tendrán valores SPAD superiores a 44 a lo largo desde los 20 a los 60 ddt aproximadamente. Mientras que las concentraciones de 0 % Fe tendrán valores SPAD inferiores a 36 hasta los 32 días ddt. A días posteriores, los valores SPAD inferiores a 36, pero superiores a 28, pertenecerán a la clase de 20 o 40 % Fe dependiendo de la ubicación en el tiempo en que se haya llevado a cabo la lectura. Dentro de los pocos trabajos similares con árboles de clasificación, Calderón *et al.* (2015) determinaron los parámetros de estrés que discriminaran entre clases de severidad de *Verticillium*. Uno de los parámetros usados fue la concentración de clorofila determinada por valores SPAD en cultivos de olivo. Bajo estas consideraciones teóricas y dado que el hierro activo o soluble (Fe^{2+}) tiene una relación con el contenido de clorofila Marsch *et al.* (1963), podría considerarse el SPAD 502 como una herramienta de detección del estatus nutrimental de Fe.

Como se observa en el Cuadro 1.1, la sensibilidad del modelo cambia de acuerdo con la concentración de Fe. El modelo tiene menos dificultad para asignar la clase al tratamiento con 0 % Fe, mientras que los tratamientos con 60 y 150 % son más difíciles de ubicar con apenas el 35 % de clasificaciones correctas. En el primer caso, hay confusión de clases con el tratamien-

to más cercano, lo cual también puede observarse con los tratamientos de 40 y 200 % Fe.

Esto podría significar un inconveniente para el desempeño del modelo debido, aunque probablemente podrían desarrollarse rangos en los que pudiera ubicarse la deficiencia del Fe.

Cuadro 1.1. Aciertos y errores de los valores observados contra valores modelados utilizando árboles de clasificación para la medición de especificidad/sensibilidad del modelo en entrenamiento.

Clase observada	Clase predicha							
	0	20	40	60	80	100	150	200
0	95	0	5	0	0	0	0	0
20	40	50	0	0	5	0	5	0
40	5	10	55	20	0	5	5	0
60	5	0	15	35	10	30	0	5
80	0	15	5	5	45	15	15	0
100	10	0	5	0	0	85	0	0
150	5	10	0	5	5	25	35	15
200	5	0	0	5	0	35	5	50
Sensibilidad (%)	95	50	55	35	45	85	35	50

Cuadro 1.2. Aciertos y errores para clase observada contra clase predicha utilizando árboles de clasificación para la medición de especificidad/sensibilidad del modelo en prueba.

Clase observada	Clase predicha							
	0	20	40	60	80	100	150	200
0	63.1	22.8	5.9	0.6	0.6	4.7	0.6	1.6
20	39.1	17.5	19.9	3.4	8.9	5.1	4.4	1.7
40	6.1	28.5	18.3	15.9	16.4	6.7	7.1	1.1
60	4.9	3.3	20.2	15.1	20.0	23.8	6.1	6.6
80	0.3	8.1	11.6	19.9	15.6	18.2	18.5	8.0
100	8.7	4.0	4.6	18.2	8.7	37.2	7.3	11.4
150	4.3	5.0	7.7	5.2	23.8	23.0	10.2	20.9
200	3.5	2.3	0.6	5.5	10.0	26.3	18.9	32.9
Sensibilidad (%)	63.1	17.5	18.3	15.1	15.6	37.2	10.2	32.9

El promedio general de clasificaciones correctas del modelo para los datos completos, es de 56.2 %, por lo que se requiere de otras variables que puedan coadyuvar a discriminar entre las diferentes clases para mejorar los porcentajes de clasificaciones correctas. Se llevó a cabo el procedimiento de validación cruzada (Cuadro 2) tomando 1000 particiones aleatorias, del total de las observaciones, aproximadamente el 88 % de los datos se usó en entrenamiento, mientras que el resto, 12 % en promedio, se utilizó en validación; en cada réplica del proceso de validación se predicen aproximadamente 20 casos. A partir de estas últimas observaciones se generó el Cuadro 1.2, donde se muestra el porcentaje de clasificaciones correctas para cada una de las clases.

Del cuadro anterior, puede notarse que el porcentaje de clasificación para el caso de 0 % Fe fue del 63 %, siendo la clase a la que se asignaron las observaciones de manera correcta. En cambio, clases como 150 % Fe tuvieron bajos porcentajes de clasificaciones correctas, ya que el algoritmo las ubicó en clases contiguas a la correcta, por lo que la sensibilidad del modelo no fue buena. Lo mismo ocurre con clases como 20 % Fe, donde la mayoría fue asignada a las clases contiguas. Probablemente, la explicación a la confusión pueda deberse más a causas fisiológicas del cultivo que a cuestiones estadísticas. En este trabajo solo se exploró la utilización de dos modelos estadísticos, sin embargo podrían probarse otros modelos como la regresión de mínimos cuadrados parciales. Lin y Huang (2011) en una investigación relacionada obtuvieron un coeficiente de correlación superior al 98% cuando se estimó el estatus del ni-

trógeno en hojas de colza, no obstante, también se utilizaron la reflectancia espectral y espectroscopía de reflectancia del infrarrojo cercano como covariables.

1.6 CONCLUSIONES

Tanto los árboles de clasificación como las redes neuronales Bayesianas regularizadas son herramientas útiles en el análisis de la información que no se ajusta a modelos lineales, ya sea para variables que puedan ser caracterizadas como cualitativas o cuantitativas. Sin embargo, con la adición de covariables podría robustecerse la modelación de la información para obtener mejores resultados en etapa de prueba del modelo debido a que en el primer caso, árboles de clasificación, el porcentaje global de clasificaciones correctas es de 44 %, mientras que en el caso de las BRNN, el coeficiente de correlación de Pearson entre valores observados y predichos puede alcanzar un valor de 0.71. No obstante, la predicción de las deficiencias de Fe en el cultivo de frijol es posible utilizando valores SPAD y su distribución en el tiempo (fenología del cultivo).

1.7 LITERATURA CITADA

- Adams, M. L., W. A. Norvell, W. D. Philpot, and J. H. Peverly. 2000. Spectral detection of micronutrient deficiency in 'Bragg' soybean. *Agronomy Journal* 92(2): 261-268.
- Ammari, T. G., A. B. Tahboub, and T. R. Abu-Zahra. 2009. Management of iron deficiency stress in citrus through soil application of vivianite to a calcareous soil. *International Journal of Botany* 5: 186-189.
- Bishop, C. M. 1995. *Neural Networks for Pattern Recognition*. Oxford University Press. 482 p.
- Bishop, C. M. and M. E. Tipping. 2003. Bayesian Regression and Classification *In: Advances in Learning Theory: Methods, Models and Applications*, J. Suykens *et al.* (eds). IOS Press, NATO Science Series III: Computer and Systems Sciences 190: 267-285.
- Bustos, M., J. R. 2005. *Inteligencia Artificial en el Sector Agropecuario*. Reporte técnico. Seminario de Investigación I: 2005-II. 299622. Universidad Nacional de Colombia, Ingeniería de Sistemas y Computación.
- Calderón, R., J. A. Navas-Cortés, and P. J. Zarco-Tejada. 2015. Early detection and quantification of verticillium wilt in olive using hyperspectral and thermal imagery over large areas. *Remote Sensing* 7: 5584-5610.
- Clark, R.B. 1991. Iron: unlocking agronomic potential. *Solutions* 35 (3):24-28.
- Correa M., C. Bielsa, J. Teixeira-Pamies, and J. R. Alique. 2008. Redes Bayesianas *vs.* redes neuronales en modelos para la predicción del acabado superficial. Technical Note IEEE Automation Society 3244-5445.
- Foresee, F. D., and M. T. Hagan. 1997. Gauss-Newton approximation to Bayesian regularization. *Proceedings of the 1997 International Joint Conference on Neural Networks*.
- Hansen, N. C., B. G. Hopkins, J. W. Ellsworth, and V. D. Jolley. 2007. Iron nutrition in field crops. *In: Iron nutrition in plants and rhizospheric microorganisms*. Barton and Abadía (eds.). Springer. 23-59.

- Hirai, M., K. Higuchi, H. Sasaki, T. Suzuki, T. Maruyama, M. Yoshiba, and T. Tadano. 2007. Contribution of iron associated with high-molecular-weight substances to the maintenance of the SPAD value of young leaves of barley under iron-deficient conditions. *Soil Science and Plant Nutrition* 53: 612–620
- James, G., D. Witten, T. Hastie, and R. Tibshirani. 2013. *An Introduction to Statistical Learning with Applications in R*. Springer Science+Business Media. 426 pp.
- Kobraee, S., K. Shamsi, and S. Ekhtiari. 2011. Soybean nodulation and chlorophyll concentration (SPAD value) affected by some of micronutrients. *Annals of Biological Research* 2 (2): 414-422.
- Köseoğlu, A. T. and V. Açıkgoza. 1995. Determination of iron chlorosis with extractable iron analysis in peach leaves. *Journal of Plant Nutrition* 18:153-161.
- Lin, S. and X. Huang. 2011. *Advances in Computer Science, Environment, Ecoinformatics, and Education*. International Conference, CSEE 2011. Proceedings, Part II. Communications in Computer and Information Science. 634 p.
- Loh, W. Y. 2011. Classification and regression trees, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1: 14–23.
- Lombard, K.; M. O’Neill, J. Mexal, A. Ulery, B. Onken, G. Bettmann, and R. Heyduck. 2010. Can soil plant analysis development values predict chlorophyll and total Fe in hybrid poplar? *Agroforestry Systems* 78: 1-11.
- López-Millán A. F., F. Morales, A. Abadía, and J. Abadía. 2001. Iron deficiency-associated changes in the composition of the leaf apoplastic fluid from field-grown pear (*Pyrus communis* L.) trees. *Journal of Experimental Botany* 52:1489–1498.
- Ma, H., H. Y. Ji, and S. L. Won. 2014. Detection of citrus greening based on Vis-NIR spectroscopy and spectral feature analysis. *Guang Pu Xue Yu Guang Pu Fen Xi* 34(10):2713-8.
- MacKay, D. J. C. 1992. Bayesian interpolation. *Neural Computation* 4 (3):415-447.

- Markwell, J., J. C. Osterman, and J. L. Mitchell. 1995. Calibration of the Minolta SPAD-502 leaf chlorophyll meter. *Photosynthesis Research* 46: 467–472.
- Marsch, H. V. Jr.; H. J. Evans, and G. Matrone. 1963. Investigation on the role of iron in chlorophyll metabolism. *Plant Physiology* 38:638-642.
- Morales, F., A. Abadía, and J. Abadía. 1991. Chlorophyll fluorescence and photon yield of oxygen evolution in iron-deficient sugar beet (*Beta vulgaris* L.) leaves. *Plant Physiology* 97:886–893.
- Nguyen, D. and B. Widrow. 1990. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *Proceedings of the IJCNN* 3: 21-26.
- Okut, H., D. Gianola, G. J. Rosa, and K. A. Weigel. 2011. Prediction of body mass index in mice using dense molecular markers and a regularized neural network. *Genetics Research* 93:1–13.
- Papasavvas, A., V. Triantafyllidis, G. Zervoudakis, G. Kapotis, Y. Samaras, and G. Salahas. 2008. Correlation of SPAD-502 meter readings with physiological parameters and leaf nitrate content in *Beta vulgaris*. *Journal of Environmental Protection and Ecology* 9 (2): 351–356.
- Pérez-Rodríguez, P., Gianola, D., Weigel, K.A., Rosa, G.J.M., and Crossa, J. 2012. Technical Note: An R package for fitting Bayesian regularized neural networks with applications in animal breeding. *Journal of Animal Science* 91:3522-3531.
- Pérez-Rodríguez, P. and D. Gianola. 2016. *brnn*: Bayesian Regularization for Feed-Forward Neural Networks. R package version 0.6. <https://CRAN.R-project.org/package=brnn>
- R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

- Razeto, B. y G. Valdés. 2006. Análisis de hierro soluble en tejidos para diagnosticar el déficit de hierro en nectarino. *Agricultura Técnica* 66(2): 216-220.
- Ripley, B. D. 1996. *Pattern recognition and neural networks*. Cambridge University Press. 416 p.
- Ripley, B. 2016. *tree: Classification and Regression Trees*. R package version 1.0-37. <https://CRAN.R-project.org/package=tree>.
- Shalizi, C. 2009. *Classification and Regression Trees*. 36-350, *Data Mining*. Carnegie Mellon University. Pittsburgh, PA.
- Soldatini, G., M. Tognini, A. Castagna, B. Baldan, and A. Ranieri. 2000. Alterations in thylakoid membrane composition induced by iron starvation in sunflower plants. *Journal of Plant Nutrition* 23:1717-1732.
- Strzepek R. F. y P. J. Harrison. 2004. Photosynthetic architecture differs in coastal and oceanic diatoms. *Nature* 431: 689–692.
- Vesali F., M. Omid, A. Kaleitab, and H. Mobli. 2014. Development of an android app to estimate chlorophyll content of corn leaves based on contact imaging. *Computers and Electronics in Agriculture* 116:211–220.
- Yadavalli, V., C. C. Jolley, C. Malleda, B. Thangaraj, P. Fromme, and R. Subramanyam. 2012. Alteration of Proteins and Pigments Influence the Function of Photosystem I under Iron Deficiency from *Chlamydomonas reinhardtii*. *PLoS ONE* 7(4): e35084.
- Zaiter, H. Z., D. P. Coyne, R. B. Clark, D. T. Lindgren, P. T. Nordquist y L. A. Pavlish. 1992. Leaf chlorosis and seed yield of dry beans grown on high-pH calcareous soil following foliar iron sprays. *Hortscience* 27(9):983-985.

C A P Í T U L O I I



DETECCIÓN TEMPRANA Y CLASIFICACIÓN DE DEFICIENCIAS DE HIERRO Y MANGANESO EN FRIJOL

Edgar García-Cruz¹, Manuel Sandoval-Villa¹, José Alfredo Carrillo-Salazar¹, Paulino Pérez-Rodríguez¹, Jorge Dionisio Etchevers-Barra¹, Antonia Macedo-Cruz¹

¹Colegio de Postgraduados, Montecillo Estado de México

2.1 RESUMEN

La identificación visual de sintomatologías en los cultivos es subjetiva y si se añade la dificultad que representa reconocer tales síntomas en condiciones de campo. Los análisis químicos tienen la función de sustentar un diagnóstico nutrimental del cultivo. Pese a que los métodos de determinación para algunos elementos se han optimizado, para algunos otros elementos, se ha dificultado establecer una correlación entre los valores obtenidos por un método de determinación y el estatus nutrimental del cultivo. Se propone entonces un método de diagnóstico nutrimental de Fe y Mn para el cultivo de frijol, no destructivo y con la identificación de patrones a partir de características de color y textura e identificados con el uso de redes neuronales artificiales. Para tal fin se establecieron ocho tratamientos que representan los niveles de Fe o Mn en la solución nutritiva (0, 50 y 200 % Fe; 0, 50 y 200 % Mn) además de una interacción (0 % Fe y 0 % Mn) en la que se suprimió el suministro de ambos micronutrientes evaluados. De forma adicional se incluyó el tratamiento testigo en la que se suministró la dosis óptima de Fe y Mn (100 % Fe y 100 % Mn). Todos los tratamientos tuvieron como base la solución universal Steiner. El muestreo se llevó a cabo en dos etapas: reproductiva de prefloración (R5) a los 63 días después del trasplante (ddt) y en la etapa reproductiva de floración (R6) a los 80 ddt. Se propone un clasificador con tres diferentes escenarios de entrada que consisten en combinaciones de características tanto de color como textura. Los porcentajes globales de clasificación indican que se pueden identificar con éxito las deficiencias de Fe y Mn en el orden de 70 % en la primera etapa, mientras que tal porcentaje aumenta hasta 80 % en la segunda etapa de muestreo.

Palabras clave: patrones; clasificador; RGB, CIELab; textura; redes neuronales.

EARLY DETECTION AND CLASSIFICATION OF IRON AND MANGANESE DEFICIENCIES IN BEAN LEAVES

Edgar García-Cruz¹, Manuel Sandoval-Villa¹, José Alfredo Carrillo-Salazar¹, Paulino Pérez-Rodríguez¹, Jorge Dionisio Etchevers-Barra¹, Antonia Macedo-Cruz¹

¹Colegio de Postgraduados, Montecillo Estado de México

2.2 SUMMARY

The visual identification of symptomatology in crops is subjective and it is difficult to recognize such symptoms in incipient conditions. Of course, the existence of chemical analysis has the function of sustaining a nutritional diagnosis of the crop. Although the methods of determination for some elements have been optimized, for some other elements, it has been difficult to establish a correlation between the values obtained by a method of determination and the nutritional status of the crop. For this research, a non-destructive nutritional diagnostic method of Fe and Mn is proposed for bean crop, by the identification of patterns based on color and texture characteristics and identified with the use of artificial neural networks. Eight treatments were established that represent Fe or Mn levels in the nutrient solution (0, 50 and 200 % Fe, 0, 50, and 200 % Mn) in addition, an interaction (0 % Fe and 0 % Mn) in which we suppressed the supply of both micronutrients evaluated. Additionally, the control treatment in which the optimum dose of Fe and Mn (100 % Fe and 100 % Mn) was supplied. All treatments were based on the universal Steiner nutrient solution. The sampling was carried out in two stages: pre-flowering reproductive (R5) at 63 days after transplant (dat) and in the reproductive stage of flowering (R6) at 80 dat. Three input scenarios with different combination of features, both color and texture, are proposed. The global percentages of classification indicate that the deficiencies of Fe and Mn can be identified successfully in the order of 70 % in the first stage, while that percentage increases up to 80 % in the second stage of sampling.

Keywords: patterns; classifiers; RGB, CIELab; texture; neural networks.

2.3 INTRODUCCIÓN

El hierro (Fe) y el manganeso (Mn) son dos de los 17 elementos esenciales que son los ladrillos usados para crear los miles de compuestos orgánicos que constituyen el tejido vegetal o llevan a cabo los procesos de crecimiento. El Fe es crucial para las plantas porque es componente de ciertas enzimas y proteínas que son usadas para la fotosíntesis (donde la energía es capturada y almacenada por las plantas) y en la respiración (donde las reservas de energía almacenadas se descomponen para liberar energía necesaria para el funcionamiento vegetal). El Mn es esencial para la fotosíntesis también; es un activador de enzimas durante la producción de clorofila y es un componente estructural de los cloroplastos donde ocurre la fotosíntesis (Zuzek y Zlesak, 2016).

Una forma de identificar la deficiencia de alguno de los dos micronutrientes mencionados es por la expresión de la sintomatología visual. Marschner (2012) afirma que los síntomas visuales son útiles para la identificación de desórdenes nutricionales, sin embargo, no son suficientes para diagnosticar el estatus nutricional de un cultivo, ya que otros factores no nutricionales, bióticos o abióticos, podrían inducir síntomas similares a aquellos producidos por las deficiencias. Además, agrega, el diagnóstico por síntomas visuales podría ser más complicado cuando más de un nutriente es deficiente o cuando la deficiencia es producida por algún otro nutriente. Por lo tanto, el diagnóstico visual es complementario y deben considerarse otras características fisiológicas o genéticas para identificar una deficiencia. Shear y Faust (1980) establecieron que los factores climáticos y de suelo afectan la absorción y función de

los nutrientes. Por lo tanto, la apariencia de un síntoma característico de la deficiencia de cierto elemento podría no indicar la falta de un elemento en el medio, o aún en la planta. La absorción del elemento podría estar inhibida por la falta o exceso de la humedad de la planta.

En el caso de la deficiencia de Fe, la sintomatología visual se manifiesta en las hojas de cítricos son similares a aquellos encontrados en otros cultivos, propiamente, nervaduras verdes bastante distinguibles del mesófilo menos verde o amarillo (Srivastava, 2012; Maldonado-Torres *et al.* 2006). El decremento en el contenido de clorofila en las hojas como respuesta a la deficiencia de Fe es probablemente el indicador más antiguo de este desorden nutricional (Gris, 1844). La clorosis de las hojas jóvenes es a menudo, el primer signo de deficiencia de Fe asociado con la pérdida de clorofila; pero también depende de cambios en la expresión y estructura de otros componentes del aparato fotosintético (Terry y Abadía, 1986); no obstante, Gonzalo *et al.* (2011) en una investigación llevada a cabo para identificar los cambios fisiológicos que se presentan en condiciones de deficiencia de Fe, encontraron que, los cambios fueron mayores en los genes involucrados en la homeóstasis del Fe que en la apariencia externa de la planta, ya que, a excepción de la clorosis, la apariencia de la planta fue similar en los materiales evaluados. La raíz no mostró un incremento de las raíces laterales, en el sistema aéreo, el crecimiento de brotes fue similar entre tratamientos. Pero, también se indica que los mecanismos de respuesta de los materiales evaluados fueron diferentes. Los genes involucrados en el metabolismo del Fe se expresaron de manera distinta, ya sea activando la quelato-reductasa férrica o los genes de transportadores de Fe o bien, la expresión de algún gen indu-

cida por la señalización a larga distancia basada en el estatus del Fe y la removilización de la planta entera.

Por otro lado, en el caso del Mn, la deficiencia de este micronutriente se puede identificar por una clorosis intervenal, mientras que las nervaduras permanecen verdes. Al igual que la deficiencia de Fe, ésta se presenta en las hojas jóvenes, pero en la deficiencia de Mn, la diferencia entre coloración de nervaduras y espacio intervenal es difusa (Hong *et al.* 2010). No obstante, Voss (1998) menciona que ambas deficiencias podrían confundirse en cultivos como la soya.

Existen métodos no paramétricos para el análisis de datos cuando estos no cumplen con las condiciones requeridas. Dentro de tales métodos, se incluyen las redes neuronales artificiales (RNA), que de acuerdo con Huang *et al.* (2010), son herramientas de modelación, usualmente de relaciones complejas entre variables de entrada y salida para encontrar patrones en datos. Su campo de aplicación es muy grande en las ciencias biológicas y agrícolas cuando se compara con otras técnicas computacionales. Por ejemplo, Vakilian y Massah (2012) diseñaron un robot con un sistema de procesamiento de imágenes en tiempo real para detectar deficiencias de nitrógeno en pepino bajo condiciones de invernadero y utilizando características texturales.

Concretamente, las RNA supervisadas con base en el algoritmo de correlación en cascada se considera adecuado para problemas de clasificación de imágenes. Existen trabajos previos, como el llevado a cabo por Dheeba y Tamil Selvi (2012) en el que, características texturales de imágenes de mamogramas con agrupaciones minúsculas de calcificaciones se usaron para detectar cáncer de mama en una etapa temprana con un porcentaje de precisión de 89 %. Los autores justifican el uso de este algoritmo al indicar que este tipo de RNA tienen la característica distintiva que no usa un conjunto predefinido de unidades ocultas, en cambio, las unidades ocultas se agregan una por una hasta minimizar el error, lo que lo hace eficiente. No obstante, el empleo de este enfoque ya se había aplicado a la investigación hecha por McKenna *et al.* (1992) con 80 características texturales extraídas de imágenes de frotis cervicales para detectar una condición, ya sea benigna o anormal como indicador de una etapa precancerígena. Dentro de las opciones para el aprovechamiento de esta herramienta de clasificación se encuentra un programa desarrollado para tal fin, NeuroShell® Classifier que, de acuerdo con la misma compañía (Ward Systems Group, Inc.), busca patrones en grupos de entrenamiento de datos, aprende los patrones y desarrolla la habilidad de clasificar correctamente nuevos patrones para hacer pronósticos o predicciones. Este software ha sido aplicado con éxito en diferentes ámbitos que involucran la modelación de reacciones químicas cinéticas y el monitoreo de la calidad de productos. En esta investigación, se propone crear un clasificador con base en las redes neuronales artificiales probabilísticas para identificar deficiencias de Fe y Mn en hojas de frijol en dos etapas fenológicas de acuerdo con la escala de Lebaron (1974): en la etapa reproductiva de prefloración (R5) a los 63 días después del transplante (ddt) cuando no eran evidentes las deficiencias parciales y en la etapa reproductiva de floración

(R6) a los 80 ddt, cuando las deficiencias fueron evidentes visualmente en los tratamientos con ausencia total de los micronutrientes Fe y Mn y síntomas incipientes en los tratamientos con deficiencias parciales.

2.4 MATERIALES Y MÉTODOS

Siembra

Se sembraron 120 semillas de frijol (*Phaseolus vulgaris* L.) cv. Cacahuete en charolas de 60 cavidades utilizando como sustrato inerte la perlita (Agrolita®), el cual es considerado inerte y estéril. Se humedeció el sustrato y posteriormente se taparon las charolas con una cubierta plástica hasta el momento de la emergencia, con la finalidad de conservar la humedad. Una vez que la radícula fue visible se retiró la cubierta y se llevaron a cabo riegos diarios; durante todo el proceso se utilizó únicamente agua destilada. Con lo anterior, se evitó la interferencia con los micronutrientes evaluados; adicionalmente se eliminaron los cotiledones a los siete después de la siembra (dds).

Trasplante y tratamiento

El trasplante se llevó a cabo 24 dds en vasos de poliestireno expandido de 1 L de capacidad. Se tomó como referencia el sistema de raíz flotante. La fertilización se hizo en el riego de acuerdo a los ocho tratamientos con las variaciones indicadas en el Cuadro 2.1, en función del micronutriente a evaluar mediante la técnica del elemento faltante. Cada tratamiento a su vez tuvo diez repeticiones. La solución nutritiva referencia fue la solución Steiner. Se usó

agua destilada para preparar la solución nutritiva. Estas soluciones nutritivas fueron ajustadas a un pH de 6.3 y CE de 2.0 dS m⁻¹ y fueron suministradas una vez que las hojas cotiledonales estaban expandidas.

Cuadro 2.1. Tratamientos evaluados en la investigación con base en la solución Steiner (1984).

Tratamiento	Concentración del micronutriente en la solución nutritiva			
	Fe (mg·L ⁻¹)	Mn (mg·L ⁻¹)	Fe (%)	Mn (%)
1 Testigo (solución Steiner)	3*	0.70*	100	100
2 Excedente en Fe	6	0.70	200	100
3 Deficiente de Fe	1.5	0.70	50	100
4 Sin Fe	0	0.70	0	100
5 Sin Mn, sin Fe	0	0	0	0
6 Sin Mn	3	0	100	0
7 Deficiente de Mn	3	0.35	100	50
8 Excedente en Mn	3	1.40	100	200

*Requerimientos para una solución Steiner completa (Alcántar y Trejo-Téllez, 2007).

Análisis de imágenes digitales y generación del clasificador con redes neuronales artificiales

Obtención de las imágenes digitales

De acuerdo con Benton (1997), la cuarta hoja recientemente madura es el órgano indicador para el diagnóstico nutrimental de Fe en el cultivo de frijol; tomando en cuenta lo establecido, se llevaron a cabo dos muestreos de la hoja antes mencionada por cada una de las repeticiones en dos momentos: en la etapa reproductiva de prefloración (R5) a los 63 ddt y en la etapa reproductiva de floración (R6) a los 80 ddt de acuerdo con la escala de Lebaron (1986).

Las hojas se almacenaron en bolsas con cierre hermético y en condiciones que evitaran la deshidratación o descomposición de las mismas. En seguida, se obtuvieron imágenes digitales de las hojas con un escáner comercial tipo “flatbed” comercial HP Scanjet G2410 y se almacenaron en el formato JPG con una resolución de 300 dpi. Concluido el escaneo de la hoja, por

cada uno de los foliolos, se tomaron seis submuestras con una dimensión de 200 por 200 píxeles. Al final, se obtuvieron 180 submuestras por cada uno de los ocho tratamientos, lo que constituye un conjunto de 1440 subimágenes. El proceso de manipulación de las imágenes hasta la obtención de las subimágenes se llevó a cabo usando el programa de distribución libre para Windows, GIMP 2.8.22 (<https://www.gimp.org>).

Análisis de color y textura

Con un programa en Microsoft® Visual Basic® se obtuvieron los valores numéricos de los píxeles para los tres canales de color del espacio RGB (Rojo, verde y azul, por sus siglas en inglés). Los valores RGB se convirtieron al clasificador de color estándar sRGB (lineales) definido por la Commission Internationale de L'Éclairage (IEC, IEC61966-2-1, 1999 citado por Mendoza *et al.* 2006), con lo que se calculó el espacio de color CIE-Lab. El croma (C) se calculó con la ecuación $(a^2+b^2)^{1/2}$, mientras que el matiz (H) se calculó a partir del arcotangente de la relación a/b (McGuire, 1992), donde a y b son dos canales del espacio de color CIE-Lab. Para el análisis de textura, las subimágenes de 200 por 200 píxeles fueron transformadas a una escala de grises de 8 bits (A). Con base en la metodología de Haralick *et al.* (1973), las imágenes A se cuantizaron a 16 tonos de grises (n). Después se obtuvo la matriz de co-ocurrencia de frecuencias relativas p_{ij} de pares de píxeles vecinos en una imagen digital, un teniendo nivel de tono i y otro j , separados por $\delta = (r = 1, \theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ)$ en la submuestra, donde r es la distancia en píxeles y θ el ángulo. $\delta = (r, \theta)$ denota un vector en las coordenadas polares de la imagen. Los valores p_{ij} por ángulo se promediaron para generar

una matriz de co-ocurrencia promedio, que fue registrada en una hoja de cálculo de n columnas y m hileras de la matriz de tonos de gris de dependencia espacial de la imagen, donde n es el número de tonos de grises (16 en este caso) y $n = m = 16$. Con los p_{ij} se obtuvieron las siguientes características texturales:

Segundo momento angular (SMA):

$$SMA = \sum_{i=1}^n \sum_{j=1}^m [p_{ij}(\delta, \theta)]^2 \quad \text{Ec. 1}$$

Entropía (EN):

$$EN = \sum_{i=1}^n \sum_{j=1}^m p_{ij}(\delta, \theta) \log[p_{ij}(\delta, \theta)] \quad \text{Ec. 2}$$

Inercia (IN):

$$IN = \sum_{i=1}^n \sum_{j=1}^m (i - j)^2 p_{ij}(\delta, \theta) \quad \text{Ec. 3}$$

Homogeneidad local (HoL):

$$HoL = \sum_{i=1}^n \sum_{j=1}^m 1/[1 + (i - j)^2] p_{ij}(\delta, \theta) \quad \text{Ec. 4}$$

Las siguientes características texturales (Ecuaciones 5 a 7) se describen con base en la metodología de Chappard *et al.* (2003). El índice de rugosidad (IR) se calculó a partir de las imágenes A . Después se obtuvieron imágenes binarias ($A\sigma$) separadas sucesivamente por umbrales definidos cada 30 niveles de grises (σ). A cada $A\sigma$ se le aplicó una “apertura morfológica” con un elemento estructurante de tres por tres que consiste de aplicar una erosión seguida de una dilación para eliminar píxeles aislados y en seguida se obtuvo su borde en la imagen bina-

ria $B^{A\sigma}$. Se contó el número de píxeles del borde $x(B^{A\sigma})$, y del área Ar a partir de la imagen A para calcular el RI:

$$RI = \left[\sum_{i=1}^n x(B^{A\sigma})_i \right] / Ar \quad \text{Ec. 5}$$

donde n es el número de imágenes binarias $A\sigma$.

Respecto a la heterogeneidad de la imagen (HE), las imágenes A se mapearon a tres niveles de grises, negro (0 a 85 se mapeó como 0), gris (86 a 171 se mapeó a 150) y blanco (172 a 255 se mapeó a 255). HE se calculó con base en las áreas negra (A_b), gris (A_g) y blanca (A_w) de cada muestra de imagen:

$$HE = (A_b + A_w) / (A_b + A_g + A_w) \quad \text{Ec. 6}$$

Para calcular la no-uniformidad del nivel de gris (GLN) se usaron las imágenes A y se cuantizaron a 16 niveles de grises. Los píxeles consecutivos del mismo valor de gris en una dirección dada constituyen un “camino” (Chu *et al.* 1990). Chappard *et al.* (2003) consideran únicamente dos direcciones, mismas que se utilizan en el presente trabajo: horizontal (GLNH) y vertical (GLNV). En cada dirección se calculó una matriz bidimensional con los elementos p_{ij} que representa el número de caminos con longitud j y de nivel de gris i . Además se calculó s que se refiere al número total de caminos en toda la imagen en una dirección determinada:

$$GLN = \sum_{i=1}^M \left(\sum_{j=1}^N p(ij) \right)^2 / s \quad \text{Ec. 7}$$

donde M es el número de niveles de gris, y N es el camino de mayor longitud en una dirección determinada para un nivel de gris.

Entrenamiento y prueba de escenarios de entradas

El programa en Visual Basic® usó una hoja de cálculo para almacenar los valores promedio por subimágen de los canales de los espacios de color anteriores. En esa misma hoja se almacenaron los valores de las características texturales. Después se usó un archivo de texto delimitado por comas para guardar los datos, con lo cual se obtuvieron 1440 datos entrada-salida. El clasificador utilizado se entrenó y probó con el sistema Neruoshell Classifier® Release 3.0 de AI Trilogy (Ward Systems Group, Inc.) con la estrategia de entrenamiento tipo neuronal supervisada, basado en el algoritmo de correlación en cascada propuesto por Fahlman y Lebiere (1990) y ejemplificado en la Figura 2.1 para uno de los escenarios de entrada propuestos.

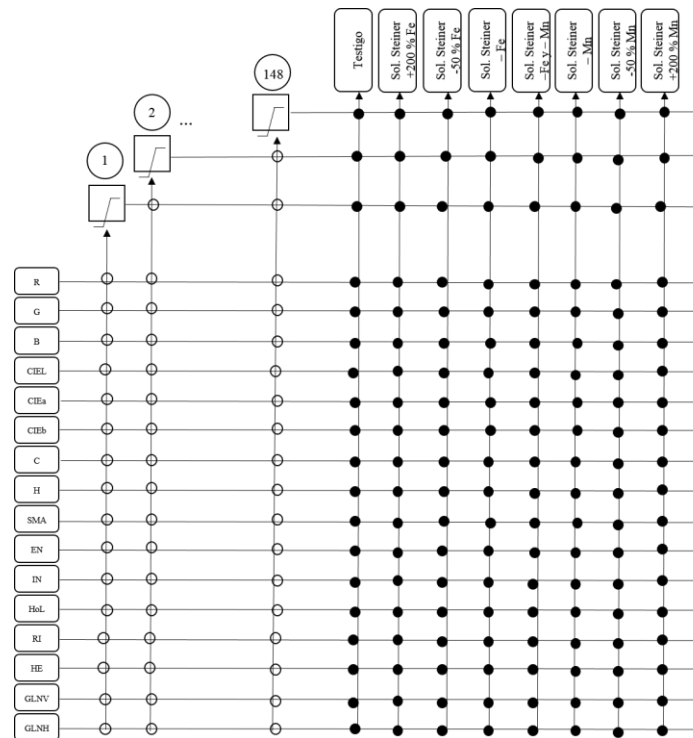


Figura 2.1. Estrategia de entrenamiento tipo neuronal supervisada; esquema basado en el algoritmo de correlación en cascada propuesto por Fahlman y Lebiere (1990) para un clasificador con 16 variables de entrada y ocho variables de salida. Los números en los círculos indican las neuronas de la capa interna que se agregan una tras otra hasta llegar a 148. Los círculos vacíos indican conexiones congeladas, mientras que los círculos negros son entrenados repetidamente.

Este algoritmo inicia con una red neuronal artificial mínima que durante el entrenamiento añade, una por una, nuevas unidades en la capa oculta, lo cual genera una estructura multicapa. Una vez que se añade a la estructura una nueva unidad en la capa oculta, los pesos del lado de las entradas se hacen constantes por lo cual esta unidad se vuelve un detector de patrones permanente en la red neuronal, y está disponible para producir valores de salida o para crear otros detectores de patrones más complejos. Esta arquitectura se caracteriza por la rapidez para entrenar las redes neuronales artificiales con pocos juegos de datos, y donde un patrón de entrada es clasificado de acuerdo con un número específico de categorías (Ward Systems Group, Inc., 1997-2007). El clasificador tuvo de 10 a 16 variables de entrada (los canales del espacio de color RGB, los espacios de color CIE-Lab, el C, y el H) y de textura (SMA, EN, IN, HoL, RI, HE, GLNH, y GLNV), ocho clases de salida o tratamientos (Cuadro 1), y un número máximo de 150 neuronas en la capa oculta. Los escenarios de entradas en los cuales el porcentaje global de clasificación correcta fuera inferior a 70 % en entrenamiento no se consideraron para evitar porcentajes inferiores en etapa de prueba. Cada escenario de entradas se repitió 10 veces con particiones aleatorias de los datos de 90 % para el entrenamiento y 10 % para la prueba. Las tablas de contingencia de uno de los mejores escenarios de entradas de cada fecha de evaluación se obtuvieron con base en la respuesta global del porcentaje de clasificación correcta promedio de las diez particiones aleatorias. Se calculó la sensibilidad, referida como la fracción de observaciones con el síntoma identificado correctamente, y la especificidad, referida como la fracción de observaciones descartadas correctamente de tener el síntoma.

2.5 RESULTADOS Y DISCUSIÓN

En esta investigación, se usaron los canales de diferentes modelos de color y características de textura obtenidas de imágenes digitales de hojas de frijol en la etapa reproductiva de preflorescencia (R5) a los 63 días después del trasplante (ddt), y en la etapa reproductiva de floración (R6) a los 80 ddt, como variables de entrada del clasificador entrenado con en redes neuronales artificiales para identificar deficiencias de hierro y manganeso.

En los mejores escenarios de entradas propuestos fue observable como se complementaron tanto variables de color como de textura para obtener un mejor desempeño evaluado mediante el porcentaje global de clasificación correcta en la prueba. En los escenarios de entrada que solo fueron consideradas las variables derivadas de los espacios de color o solo las características de textura, el desempeño fue considerablemente menor (resultados no presentados). Choras (2007) menciona que las características de textura podrían ayudar a diferenciar imágenes con textura de las que carecen de ella, y que normalmente la visión humana es incapaz de detectar adecuadamente. Por otro lado, el color hace más efectivo el proceso de clasificación de píxeles cuando se emplean características de textura (Anami *et al.* 2010).

Los valores de desempeño de los escenarios de entradas propuestos se presentan en los Cuadros 2.2 y 2.4. En la primera fecha de muestreo, el rango de porcentaje global de clasificación correcta en la prueba fue entre 69.4 y 70.1 %.

Cuadro 2.2. Entrenamiento y prueba de escenarios de entradas para crear el clasificador con redes neuronales artificiales e identificar ocho clases de salida o niveles de fertilización con base en la textura y color de hojas de frijol en el primer muestreo (63 ddt): (1) Solución Steiner, Testigo; (2) Sol. Steiner +200 % Fe; (3) Sol. Steiner -50 % Fe; (4) Sol. Steiner -Fe; (5) Sol. Steiner -Fe y - Mn; (6) Sol. Steiner -Mn; (7) Sol. Steiner -50 % Mn y (8) Sol. Steiner +200 % Mn.

#	Variabes incluidas en el escenario de entrada	CCE (%)	NN	CCP (%)	Dstd
1	R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH	78.7	148	70.1	3.6
2	R, G, B, CIEL, CIEa, CIEb, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH	74.7	146	69.5	3.3
3	R, G, B, CIEL, CIEa, CIEb, C, H, RI, HE, GLNV, GLNH	75.2	143	69.4	3.1

CCE: Clasificación correcta en el entrenamiento; NN: Número de neuronas de la capa interna; CCP: Clasificación correcta en la prueba; Dstd: Desviación estándar de la CCP; R: Rojo, G: Verde, y B: Azul son canales del espacio de color RGB; CIEL: Luminosidad, CIEa y CIEb: Canales de color del espacio de color CIELab; C: cromia; H: matiz; SMA: Segundo momento angular; EN: Entropía; IN: Inercia; HoL: Homogeneidad local; RI: Índice de rugosidad; HE: Heterogeneidad; GLNV: No uniformidad del nivel de gris en dirección vertical; GLNH: No uniformidad del nivel de gris en dirección horizontal.

Cuadro 2.3. Estadísticas promedio de 10 arreglos matriciales (tablas de contingencia) para el escenario de entradas: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la prueba, y ocho clases de salida: (1) Solución Steiner, Testigo; (2) Sol. Steiner +200% Fe; (3) Sol. Steiner -50% Fe; (4) Sol. Steiner - Fe; (5) Sol. Steiner -Fe y - Mn; (6) Sol. Steiner - Mn; (7) Sol. Steiner -50% Mn y (8) Sol. Steiner +200% Mn. Muestreo a los 63 ddt.

		Clase observada								Total	Valor predictivo positivo (%)
		1	2	3	4	5	6	7	8		
Clase calculada	1	9.0	1.6	0.9	0.4	0.0	0.0	1.2	0.6	13.7	65.3
	2	2.0	10.0	1.0	0.0	0.0	0.2	2.4	0.8	16.4	60.9
	3	2.7	1.2	12.6	0.0	0.0	0.9	0.5	0.2	18.1	70.3
	4	0.7	0.1	0.2	12.6	0.2	1.1	0.2	1.1	16.2	78.4
	5	0.0	0.0	0.0	0.9	18.4	0.7	0.0	0.0	20.0	91.9
	6	0.0	0.1	0.9	2.1	1.0	16.6	0.1	0.1	20.9	78.9
	7	2.8	2.1	0.2	0.5	0.1	0.1	10.6	4.5	20.9	51.4
	8	1.2	0.5	0.4	0.6	0.0	0.4	3.0	9.8	15.9	63.7
Total		18.4	15.6	16.2	17.1	19.7	20.0	18.0	17.1	142	
		0.49	0.65	0.77	0.73	0.94	0.82	0.59	0.57		Proporción verd.-pos.
		0.04	0.05	0.04	0.03	0.01	0.04	0.08	0.05		Proporción falso-pos.
		0.96	0.95	0.96	0.97	0.99	0.96	0.92	0.95		Proporción verd.-neg.
		0.51	0.35	0.23	0.27	0.06	0.18	0.41	0.43		Proporción falso-neg.
		0.49	0.65	0.77	0.73	0.94	0.82	0.59	0.57		Sensibilidad
		0.96	0.95	0.96	0.97	0.99	0.96	0.92	0.95		Especificidad

En el Cuadro 2.3 se observan diferencias en la sensibilidad entre las clases de salida del escenario de entrada 1 (Cuadro 2.2). Las clases de salida con porcentajes de clasificación correcta, proporciones verdadero-positivas o sensibilidad superiores a 0.7 son aquellas que corresponden a los tratamientos en los que no se aplicó Fe o Mn (clases de salida 4 a 6, Cuadro 2.3) y en el que se aplicó 50 % de Fe (Clase 3). En esta investigación, la sensibilidad es la capacidad del clasificador para identificar correctamente un nivel de fertilización con base en variables de color y textura de segmentos de imágenes de hoja.

Cuadro 2.4. Entrenamiento y prueba de escenarios de entradas para crear el clasificador con redes neuronales artificiales e identificar ocho clases de salida o niveles de fertilización con base en la textura y color de hojas de frijol en el segundo muestreo (80 ddt): (1) Solución Steiner, Testigo; (2) Sol. Steiner +200 % Fe; (3) Sol. Steiner -50 % Fe; (4) Sol. Steiner -Fe; (5) Sol. Steiner -Fe y - Mn; (6) Sol. Steiner -Mn; (7) Sol. Steiner -50 % Mn y (8) Sol. Steiner +200 % Mn.

#	Variables incluidas en el escenario de entrada	CCE (%)	NN	CCP (%)	Dstd
1	R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH	86.6	141.7	79.5	3.7
2	R, G, B, CIEL, CIEa, CIEb, RI, HE, GLNV, GLNH	82.4	141.2	78.5	3.2
3	R, G, B, CIEL, CIEa, CIEb, C, H, RI, HE, GLNV, GLNH	83.9	144.5	80.2	3.5

CCE: Clasificación correcta en el entrenamiento; NN: Número de neuronas de la capa interna; CCP: Clasificación correcta en la prueba; Dstd: Desviación estándar de la CCP; R: Rojo, G: Verde, y B: Azul son canales del espacio de color RGB; CIEL: Luminosidad, CIEa y CIEb: Canales de color del espacio de color CIELab; C: croma; H: matiz; SMA: Segundo momento angular; EN: Entropía; IN: Inercia; HoL: Homogeneidad local; RI: Índice de rugosidad; HE: Heterogeneidad; GLNV: No uniformidad del nivel de gris en dirección vertical; GLNH: No uniformidad del nivel de gris en dirección horizontal.

La clase de salida con deficiencia de 50 % de Mn (clase 7), aquellas clases de salida con excesos nutrimentales (clases 2 y 8), y el testigo (clase 1) presentaron sensibilidades inferiores a 0.7. El testigo fue el tratamiento que presentó la menor sensibilidad (0.49) con niveles de confusión en la asignación correcta del tratamiento entre 10.8 a 15.2 % con cada una de las clases de salida 2, 3, y 7; y de 6.5 % con la clase de salida 8, lo cual implica dificultad para detectar si, a partir de imágenes digitales de hojas, el cultivo ha sido correctamente fertilizado con la solución Steiner completa. Cada una de las clases 2, 3 y 7 presentaron niveles de confusión

con el testigo de 5.5 a 10.2 %. Por otro lado, aquellas hojas provenientes del tratamiento con la solución completa también tienen un grado de confusión con el tratamiento con un exceso de Fe, y en menor grado con la condición de exceso de Mn. Los niveles de confusión para este caso, son iguales o inferiores al 2 %.

Cuadro 2.5. Estadísticas de 10 arreglos matriciales (tablas de contingencia) para la combinación de las variables de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH y ocho clases de salida: (1) Solución Steiner, Testigo; (2) Sol. Steiner +200% Fe; (3) Sol. Steiner -50% Fe; (4) Sol. Steiner – Fe; (5) Sol. Steiner –Fe y – Mn; (6) Sol. Steiner – Mn; (7) Sol. Steiner -50% Mn y (8) Sol. Steiner +200% Mn. Muestreo a los 80 ddt.

		Clase verdadera								Total	Valor predictivo positivo (%)
		1	2	3	4	5	6	7	8		
Clase predicha	1	11.1	2.3	1.8	0	0.1	0	0	0.4	15.7	69.9
	2	2.9	13.1	2	0.1	0	0	0	0	18.1	72.6
	3	2.9	2	9.3	0	0	0	0.2	0	14.4	64.4
	4	0	0	0.2	16.8	0	0.1	0	0.4	17.5	95.9
	5	0	0	0	0	13.4	0	0	0	13.4	100.0
	6	0.1	0.4	0	0.5	0	17	0.4	1.7	20.1	84.4
	7	0.4	0.3	0.3	0	0.1	0.2	11.8	2.9	16	71.9
	8	0.2	0.2	0	0.1	0	1.3	2.9	15.1	19.8	76.9
Total		17.6	18.3	13.6	17.5	13.6	18.6	15.3	20.5	135	
		0.62	0.72	0.69	0.96	0.99	0.91	0.75	0.74		Proporción verd.-pos.
		0.04	0.04	0.04	0.01	0.00	0.03	0.03	0.04		Proporción falso-pos.
		0.96	0.96	0.96	0.99	1.00	0.97	0.97	0.96		Proporción verd.-neg.
		0.38	0.28	0.31	0.04	0.01	0.09	0.25	0.26		Proporción falso-neg.
		0.62	0.72	0.69	0.96	0.99	0.91	0.75	0.74		Sensibilidad
		0.96	0.96	0.96	0.99	1.00	0.97	0.97	0.96		Especificidad

La tendencia de confusión en general no se mantiene a lo largo del tiempo como puede observarse arriba (Cuadro 2.5). Puede destacarse entonces que, para la etapa en la que se llevó a cabo el muestreo, la identificación de deficiencias de Fe cuando se usa el clasificador propuesto es adecuada; mientras que la identificación del tratamiento al que se le suministró la solución completa es deficiente, pese a que este tratamiento es confundido con aquél al que se le

suministró el 50 % de la solución completa ya sea de Fe o Mn. Esto significa que el umbral entre las características de color y textura de una hoja sana y una con deficiencia de alguno de los dos micronutrientes mencionados es cercano. En cuanto a las clases de salida 2 y 3, presentaron niveles de confusión mutuo del orden de 6.1 a 7.6 %, mientras que entre las clases de salida 7 y 8 el nivel de confusión mutuo fue de 16.6 a 26.3 %; estos tratamientos corresponden a la solución Steiner -50 % Mn y la solución Steiner +200 % Mn, respectivamente. Es notable también 13 % de confusión mutuo entre las clases 2 y 7. Respecto a las proporciones verdadero-positivas y verdadero-negativas, a medida que sus valores son más cercanos a la unidad, el desempeño del clasificador es mejor, mientras que si sus valores son más cercanos a cero, el desempeño del clasificador es insuficiente. En el primer caso, tales proporciones son un mejor indicador para aquellos tratamientos en los que se indujo algún tipo de deficiencia, a excepción de la deficiencia de Mn. Tales valores indican la proporción de submuestras que se asignaron a un tratamiento cuando verdaderamente correspondían a esa clase, por lo que está estrechamente ligada con la sensibilidad. Mientras que la proporción verdadero-negativa indica la capacidad del clasificador para descartar una submuestra de la categoría que no le corresponde. El desempeño del clasificador es superior al 90 % en todos los casos. Por otro lado, se observa que la especificidad del modelo fue bastante alta (Cuadro 2.3) con porcentajes superiores al 90 %. A partir de estos valores, puede señalarse que la eficacia del modelo para descartar clases que no corresponden a la indicada es adecuada y esto asociado a la proporción verdadero-negativo. Las proporciones falso-positivo y falso-negativo son lo opuesto y se refieren a aquellas clases que se asignaron a un tratamiento al cual no pertenecen, por lo que los valores cercanos a cero son deseables.

	Hoja muestra	Foliolo izquierdo	Foliolo central	Foliolo derecho
Testigo				
Sol. Steiner +200 % Fe				
Sol. Steiner -50 % Fe				
Sol. Steiner -Fe				
Sol. Steiner -Fe y -Mn				
Sol. Steiner -Mn				
Sol. Steiner -50 % Mn				
Sol. Steiner +200 % Mn				

Figura 2.2. Cuarta hoja de frijol var. Cacahuete, recientemente madura, de una repetición de cada uno de los tratamientos indicados. Las hojas se recolectaron a los 63 días después de la siembra.

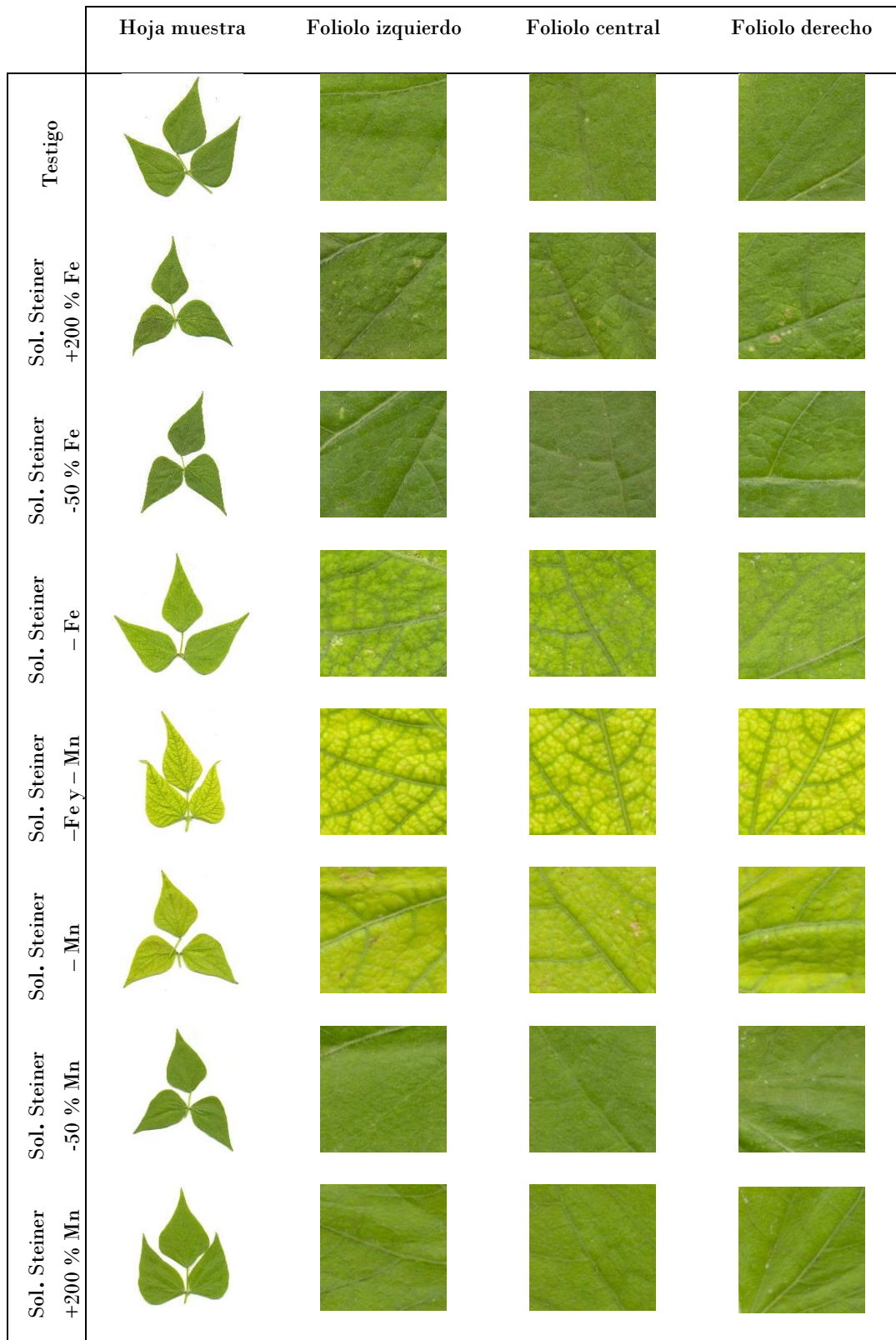


Figura 2.3. Cuarta hoja de frijol var. Cacahuate, recientemente madura, de una repetición de cada uno de los tratamientos indicados. Las hojas se recolectaron a los 80 días después de la siembra.

En la segunda fecha de muestreo, cuando los síntomas de las deficiencias totales y parciales se habían intensificado (17 días después del primer muestreo) y usando dos escenarios de entrada iguales al caso anterior, se obtuvieron porcentajes globales de clasificación correcta de alrededor de 80 %, lo que significó un aumento de 10 % con respecto a la primera fecha de muestreo (Cuadros 2.2 y 2.4). Los síntomas de los tratamientos fueron más evidentes por la coloración y textura, lo cual se observa si se comparan las Figuras 2.1 y 2.2.

En el Cuadro 2.5 se observa que mejoró la sensibilidad de las clases de salida o tratamientos en la segunda fecha de muestreo, donde a excepción del tratamiento con la solución Steiner con -50 % Fe (clase 3) y el testigo, las clases presentaron sensibilidades mayores a 70 %. De hecho, la sensibilidad de la clase 3 fue la única que descendió en el muestreo de la segunda fecha. En particular, la sensibilidad fue mayor a 90 % en los tratamientos en donde no se aplicó Fe o Mn (clases de salida 4 a 6, Cuadro 2.5) y el testigo nuevamente presentó la menor sensibilidad (0.62) y se confundió con las clases 2, y 3 en 16 %, mientras que la confusión mutua entre las clases 7 y 8 disminuyó a un rango de 14.1 a 18.9 % (Cuadro 2.5). En cambio, la confusión mutua entre las clases 2 y 3 aumentó a un rango de 10.9 a 14.7 %, y estas clases se confundieron en alrededor de 13 % con el testigo. Lo que es notable es la confusión mutua de 7 a 8 % entre las clases 6 y 8 que corresponden a los tratamientos sin y con exceso de Mn. De acuerdo con los procesos fisiológicos que ocurren en una planta bajo esta condición de deficiencia, Shi *et al.* (2012) y Millaleo *et al.* (2010) consideran que la deficiencia de Mn ocurre como un desorden latente con síntomas visuales no aparentes (Figuras 2.1 y 2.2), por lo que la

magnitud de la deficiencia de este micronutriente, constituye un problema mayor del que se esperaría cuando los síntomas visuales aparecen debido al tiempo que transcurre entre el inicio y la detección de esta deficiencia.

Si se observa en los Cuadros 2.3 y 2.5, el porcentaje de confusión entre el testigo y la clase bajo condiciones de exceso de Mn es baja (6.5 y 1.1 %) para las dos fechas de muestreo respectivamente, lo que significa una ventaja del escenario de entrada propuesto, ya que pueden ser diferenciadas ambas condiciones aunque visualmente se aprecian coloraciones parecidas. Valdés-López *et al.* (2010) en estudios previos reportaron que para inducir una toxicidad por Mn en *Phaseolus vulgaris* sometieron las plantas a una concentración 200 μM de MnCl_2 ; en esa solución, la proporción de Mn corresponde a 10.98 $\text{mg}\cdot\text{L}^{-1}$ mientras que la usada para generar la condición de exceso en esta investigación, es apenas 0.6 $\text{mg}\cdot\text{L}^{-1}$.

2.6 CONCLUSIONES

Bajo diferentes condiciones de deficiencias nutrimentales de Fe o Mn en el cultivo de frijol, tanto la textura como el color de las hojas forman patrones específicos que pueden ser clasificados usando redes neuronales artificiales con el propósito de detectar tales deficiencias antes que puedan ser diagnosticadas visualmente. El desempeño de los escenarios de entrada propuestos es adecuado en una etapa en la que, con el diagnóstico visual tradicional, las deficiencias de los micronutrientes Fe y Mn no serían detectables dado que serían incipientes. El desempeño del modelo mejora considerablemente con el paso del tiempo, aún considerando

que, niveles de deficiencia de hasta el 50 % de los dos micronutrientes estudiados, aún no producen síntomas visuales evidentes. Puede considerarse un método alternativo a los análisis de laboratorio, especialmente cuando se requiere de tomar decisiones respecto a la nutrición del cultivo.

2.7 LITERATURA CITADA

- Anami, B. S., S. S. Nandyal, and A. Govardhan. 2010. A combined color, texture and edge features based approach for identification and classification of indian medicinal plants. *International Journal of Computer Applications* 6(12): 45-51 pp.
- Abadía, J. 1994. Physiology and biochemistry of leaves under iron deficiency in iron nutrition in soils and plants. *Proceedings of the Seventh International Symposium on Iron Nutrition and Interactions in Plants*. Springer Science Business Media.
- Alcántar G., G. y Trejo-Téllez, L. I. 2007. *Nutrición de Cultivos*. Colegio de Postgraduados. 451 pp.
- Benton Jones, J. Jr. 1997. *Plant Nutrition Manual*. CRC Press. 230 pp.
- Cadena, J. A., Pérez-Londoño y S. M.; Cadena, J. M. 2008. Aplicación de redes neuronales probabilísticas en la detección de fallas incipientes en transformadores. *Scientia et Technica XIV* (39) 48–53.
- Chappard, D., I. Degasnea, G. Huré, E. Legrand, M. Audran, and M.F. Baslé. 2003. Image analysis measurements of roughness by texture and fractal analysis correlate with contact profilometry. *Biomaterials* 24: 1399–1407.
- Chu A, C. M. Sehgal, and J. F. Greenleaf. 1990. Use of gray value distribution of run lengths for texture analysis. *Pattern Recognition Letters* 11:415–9.
- Fahlman, S. E. and C. Lebiere. 1990. The Cascade-Correlation learning architecture. CMU-CS-90-100. School of Computer Science. Carnegie Mellon University. Pittsburgh, PA, USA.
- Gonzalo, M. J., M. A. Moreno, and Y. Gogorcena. 2011. Physiological responses and differential gene expression in *Prunus* rootstocks under iron deficiency conditions. *Journal of Plant Physiology* 168:887–893.

- Gris, E. 1844. Nouvelles expériences sur l'action des composés ferrugineux solubles, appliqués à la végétation, et spécialement au traitement de la chlorose et de la débilité des plantes. *Comptes Rendus de l'Académie des Sciences* 19:1118–1119.
- Haralick, R. M., K. Shanmugan, and I. Dinstein. 1973. Textural features for image classification. *IEEE Trans. Systems, Man and Cybernetics* 3(6): 610-621.
- Hong, E., Q. Ketterings, M. McBride. 2010. Manganese. Fact Sheet 49. Agronomy Fact Sheet Series. Nutrient Management Spear Program. Cornell University Cooperative Extension.
- Lebaron, M. J. 1974. Development stages of the common bean plant. University of Idaho, Cooperative Extension Service. Current information series No. 228. 2 pp.
- Maldonado-Torres R., J. D. Etchevers-Barra, and G. Alcántar-González. 2006. Morphological changes in leaves of Mexican lime affected by iron chlorosis. *Journal of Plant Nutrition*. 29:615–628.
- Marschner, P. 2012. *Marschner's Mineral Nutrition of Higher Plants*. London: Academic Press.
- Matthew L. Adams, Wendell A. Norvell, William D. Philpot, and John H. Peverly. 2000. Toward the discrimination of manganese, zinc, copper, and iron deficiency in 'Bragg' soybean using spectral detection methods. *Agronomy Journal* 92:268–274.
- McKenna S J., I.W. Ricketts, A.Y. Cairns, and K.A. Hussein. 1992. Cascade-correlation neural networks for the classification of cervical cells. *Neural Networks for Image Processing Applications*. IEE Colloquium. pp 5/1-5/4.
- Mendoza, F., P. Dejmeck, and L. Aguilera. 2006. Calibrated color measurements of agricultural foods using image analysis, *Postharvest Biology and Technology* 41 (3): 285-295.

- Millaleo, R.; M. Reyes-Díaz, A.G. Ivanov, M.L. Mora, and M. Alberdi. 2010. Manganese as essential and toxic element for plants: transport, accumulation and resistance mechanisms. *Journal of Soil Science and Plant Nutrition* 10 (4): 476 - 494
- Shear, C. B. and M. Faust. 1980. Nutritional Ranges in Deciduous Tree Fruits and Nuts. *Horticultural Reviews*, Volume 2.
- Shi, L., M. Hall, C. Funk, and W. P. Schröder. 2012. Review: Photosystem II, a growing complex: Updates on newly discovered components and low molecular mass proteins. *Biochimica et Biophysica Acta* 1817:13–25.
- Srivastava, A. K. 2012. Integrated nutrient management in *Advances in Citrus Nutrition*. Springer Verlag, p 369 – 389.
- Terry N., J. Abadía. 1986. Function of iron in chloroplasts. *Journal of Plant Nutrition* 9: 609–646.
- Valdés-López, O.; S. S. Yang, R. Aparicio-Fabre, P. H. Graham, J. L. Reyes, C. P. Vance, and G. Hernández. 2010. MicroRNA expression profile in common bean (*Phaseolus vulgaris*) under nutrient deficiency stresses and manganese toxicity. *New Phytologist* 187: 805–818.
- Vakilian, K. A. and J. Massah. 2012. Design, development and performance evaluation of a robot to early detection of nitrogen deficiency in greenhouse cucumber (*Cucumis sativus*) with machine vision. *International Journal of Agricultural Research*. 2 (4): 448-454.
- Voss, R. 1998. Micronutrients. Department of Agronomy. Iowa State University. (http://www.agronext.iastate.edu/soilfertility/info/Micronutrients_VossArticle.pdf).
- Zuzek, K. and D. Zlesak. 2016. Iron and manganese deficiencies in woody plants. University of Minnesota, Extension (<http://www.extension.umn.edu/garden/yard-garden/trees-shrubs/iron-chlorosis/>).

C A P Í T U L O I I I



INTERACCIÓN DE DEFICIENCIAS DE HIERRO Y NITRÓGENO Y SU IDENTIFICACIÓN CON REDES NEURONALES ARTIFICIALES EN FRIJOL

Edgar García-Cruz¹, Manuel Sandoval-Villa¹, José Alfredo Carrillo-Salazar¹, Paulino Pérez-Rodríguez¹, Jorge Dionisio Etchevers-Barra¹, Antonia Macedo-Cruz¹

¹Colegio de Postgraduados, Montecillo Estado de México

3.1 RESUMEN

Se ha demostrado la relación directa entre el nitrógeno (N) y la clorosis causada por la falta de este elemento, sin que sea la única causa de este desorden fisiológico. No obstante, se han propuesto modelos con los que se han identificado adecuadamente las deficiencias de N en diversos cultivos. De forma independiente, también se ha encontrado que pueden identificarse las deficiencias de hierro (Fe) utilizando la coloración de las hojas. Sin embargo, se requiere conocer el efecto de la interacción de ambas deficiencias en los patrones que se forman las características de color y textura en las hojas del cultivo de frijol, por lo que se retoma tal objetivo en esta investigación. Se plantean nueve condiciones de nutrición con base en la solución Steiner (1) -N, -Fe; (2) -N, -50 % Fe; (3) -N; (4) -50 % N, -Fe; (5) -50 % N, -50 % Fe; (6) -50 % N; (7) -Fe; (8) -50 % Fe y (9) solución Steiner completa como testigo. Se proponen, además, tres escenarios de entrada con el uso de redes neuronales artificiales mediante el software Neuroshell Classifier® para la identificación de deficiencias de Fe y N en cuatro etapas de cultivo (35, 56, 67 y 78 días después del trasplante) y tomando una hoja recientemente madura para digitalizarse. La precisión del clasificador es de 63 a 66 % para identificar las deficiencias de Fe o N y sus interacciones cuando se consideraron las cuatro fechas de muestreo en conjunto, y hasta 76 % cuando se consideran los muestreos de manera separada. La sensibilidad de para el primer caso fue 66 % y 77 % en promedio para fechas de muestreo separadas; mientras que la especificidad fue en el orden de 96 y 97 % para ambos casos.

Palabras clave: clorosis; clasificador; relación Fe-N; características texturales; color.

INTERACTION OF IRON AND NITROGEN DEFICIENCIES AND ITS IDENTIFICATION USING ARTIFICIAL NEURAL NETWORKS IN BEAN LEAVES

Edgar García-Cruz¹, Manuel Sandoval-Villa¹, José Alfredo Carrillo-Salazar¹, Paulino Pérez-Rodríguez¹, Jorge Dionisio Etchevers-Barra¹, Antonia Macedo-Cruz¹

¹Colegio de Postgraduados, Montecillo Estado de México

3.2 SUMMARY

The direct relationship between nitrogen (N) and chlorosis caused by the lack of this element has been demonstrated in previous works, however, this is not the only cause of that physiological disorder. Still, models have been proposed to identify the deficiencies of N in different crops adequately. Independently, it has also been found that the deficiencies of iron (Fe) can be identified using the coloring of the leaves. However, it is required to know the effect of the interaction of both deficiencies in the patterns that form the characteristics of color and texture in the leaves of the bean crop; for that reason, this objective is addressed in this investigation. Nine conditions of nutrition are proposed based on the solution Steiner (1) -N, -Fe; (2) -N, -50% Fe; (3) -N; (4) -50 % N, -Fe; (5) -50 % N, -50 % Fe; (6) -50 % N; (7) -Fe; (8) -50 % Fe and (9) complete Steiner solution as a control. In addition, three input scenarios with the use of artificial neural networks are proposed using the Neuroshell Classifier® software to identify deficiencies of Fe and N in four stages of cultivation (35, 56, 67 and 78 days after transplant) and taking a recently mature leaf to be digitized. The range precision of the classifier is 63 to 66% to identify the deficiencies of Fe or N and their interactions when considering the four sampling dates as a whole, and up to 76 % when the samples are considered separately. The sensitivity for the first case was 66 % and 77 % on average for separate sampling dates; while the specificity was in the order of 96 and 97 % for both cases.

Keywords: chlorosis; classifiers; Fe-N ratio; textural characteristics; color.

3.3 INTRODUCCIÓN

La relación entre el contenido de clorofila y el estatus nutrimental del nitrógeno (N) ha sido demostrada en numerosas ocasiones dado el papel principal del N en la conformación de la molécula de este compuesto; Kergoat *et al.* (2008) concluyeron que el contenido de N en el cultivo es un factor fundamental involucrado en la determinación de la eficiencia óptima del uso de la radiación en la banda fotosintética, así como la eficiencia del follaje. De manera particular en cultivos como el maíz (Ercoli *et al.* 1993, Gitelson *et al.* 2003, Ciganda *et al.* 2009, Schlemmer *et al.* 2013) se ha evaluado la correlación del contenido de clorofila de las hojas en estas plantas con el contenido de N de las mismas. Otras investigaciones han sido llevadas a cabo partiendo de lo dicho anteriormente, con lo que se han desarrollado métodos altamente eficientes en la estimación del estatus nutrimental de N en varios cultivos, ya sea con el uso de instrumentos como el SPAD 502 (Minolta Corp.) (Argenta *et al.* 2004) o bien con la propuesta de modelos de estimación de N en varias especies (Kalacska *et al.* 2015).

Por otro lado, se han llevado a cabo investigaciones en las que se ha demostrado que los análisis de Fe total no son el mejor indicador del estatus nutrimental de Fe (Abadía *et al.* 2000). Sin embargo, se ha encontrado que una fracción de este elemento sí guarda una relación cercana con el diagnóstico de la nutrición de Fe en cultivos como el durazno. Razeto y Valdéz (2006) probaron la fracción total y la soluble (o activo) como indicadores del suministro de Fe en nectarinas (*Prunus persica*) a través de la determinación de dichas fracciones en los tejidos secos de las hojas y brotes, de forma similar a como se llevan a cabo los análisis de tejido ve-

getal en un laboratorio convencional. Concluyeron que la concentración de Fe total en la hoja no representó el estatus nutrimental de Fe, mientras que la fracción soluble tuvo mejor correlación. A su vez, encontraron también que tal fracción está directamente relacionada con el contenido de clorofila en el cultivo en el orden del 83 % cuando se llevó a cabo la determinación en los brotes.

Por su parte, Köseoğlu y Açikgöz (1995) encontraron una correlación alta entre la concentración de clorofila y la de Fe soluble en la hoja de *P. persica*, tanto en tejido fresco como en tejido seco en el orden de 0.93. Sus resultados confirmaron que la concentración de Fe total no está asociada a la clorosis férrica. Sin embargo, también mencionan la dificultad para la determinación de esta fracción de Fe dada la naturaleza del tejido. Aunado a esto, Hochmuth (2011) menciona que los análisis de tejido vegetal son de difícil interpretación dado que, por ejemplo, el polvo es una fuente de contaminación y que dentro de sus componentes se encuentra también el Fe. Por lo tanto, encontrar algún método alternativo a los propuestos significaría una ventaja en la determinación del estatus nutrimental de algún cultivo.

Si bien, la deficiencia de Fe y N no podrían ser confundidas ya que los síntomas ocurren en partes distintas de la planta se requiere evaluar la interacción. La deficiencia de Fe ocurre como clorosis de las hojas más jóvenes; en estas hojas a menudo se presenta de forma interveinal, mientras que las nervaduras permanecen verdes por un tiempo mientras la deficiencia continúa. Por su lado, la deficiencia de N se manifiesta primero en las hojas más viejas, y es

en esta parte dónde se observa la clorosis (Hochmuth, 2011). No obstante se ha demostrado el rol de ambos elementos en el proceso de la fotosíntesis y las variaciones en la coloración.

En cultivos como el arroz, Chen *et al.* (2014) propusieron ya un método para identificar deficiencias de N, fósforo (P) y potasio (K) utilizando la tecnología de escaneo y considerando características de color y forma de la hoja con porcentajes de éxito de 86 a 89 %.

Con base en los antecedentes descritos, el objetivo de la presente investigación es proponer un clasificador con redes neuronales artificiales (RNA) para identificar deficiencias de Fe y de N en cuatro etapas de desarrollo del cultivo de frijol con base en características texturales y variables de color obtenidas mediante análisis de imágenes.

3.4 MATERIALES Y MÉTODOS

Siembra

El experimento se estableció en Texcoco, México. Se sembraron 100 semillas de frijol (*Phaseolus vulgaris* L.) var. Negro Chapingo en charolas de poliestireno y como sustrato, perlita (Agrolita®), que se considera inerte y estéril. Posteriormente se humedeció el sustrato y se taparon las charolas con una cubierta plástica negra hasta el momento de la emergencia (3 d), con la finalidad de conservar la humedad y temperatura. Una vez que la radícula fue visible se retiró la cubierta y se llevaron a cabo riegos diarios; durante todo el proceso se utilizó únicamente agua destilada. Con lo anterior, se evitó la interferencia con los micronutrientos

evaluados; adicionalmente se eliminaron los cotiledones a los siete días después de la siembra (dds).

Trasplante y tratamiento

El trasplante se llevó a cabo 15 dds en vasos de poliestireno expandido de 1 L de capacidad. Se tomó como referencia el sistema de raíz flotante. La fertilización se hizo en el riego de acuerdo a los nueve tratamientos con las variaciones indicadas en el Cuadro 3.1, en función del micronutriente a evaluar mediante la técnica del elemento faltante. Cada tratamiento a su vez tuvo cinco repeticiones. La solución nutritiva de referencia fue la solución Steiner. Se usó agua destilada para preparar la solución nutritiva. Los tratamientos fueron suministradas una vez que las hojas cotiledonales estaban expandidas.

Cuadro 3.1. Tratamientos aplicados con base en la solución Steiner (1984) a las plantas de frijol var. Negro Chapingo a los 15 dds.

Tratamiento	Concentración en la solución nutritiva			
	N (meq·L ⁻¹)	Fe (mg·L ⁻¹)	N (%)	Fe (%)
1 Deficiencia total Fe y N	0	0	0	0
2 Deficiencia parcial de Fe y total de N	0	1.5	0	50
3 Deficiencia total de N	0	3	0	100
4 Deficiencia total de Fe y parcial de N	6	0	50	0
5 Deficiencia parcial de Fe y N	6	1.5	50	50
6 Deficiencia parcial de N	6	3	50	100
7 Deficiencia total de Fe	12	0	100	0
8 Deficiencia parcial de Fe	12	1.5	100	50
9 Testigo*	12	3	100	100

*Requerimientos para una solución Steiner completa (Alcántar y Trejo-Téllez, 2007).

Análisis de imágenes digitales y generación del clasificador con redes neuronales artificiales

Obtención de las imágenes digitales

Las imágenes digitales se obtuvieron de las hojas de plantas de frijol a las cuales se les impusieron los tratamientos especificados en el Cuadro 3.1. Estas hojas se almacenaron en bolsas plásticas con cierre hermético y en condiciones que evitaran la deshidratación. Para efectuar un análisis químico, Benton (1997) sugiere tomar la cuarta hoja recientemente madura de una planta. Esta hoja se considera el órgano indicador para el diagnóstico nutrimental de Fe en frijol por lo que esta misma hoja se muestreó para el caso de esta investigación. Tal hoja de diferentes muestras se tomó en cuatro etapas de desarrollo del cultivo de frijol: 35, 56, 67 y 78 ddt que corresponden con las etapas fenológicas respectivamente descritas en la escala de Lebaron (1986). Se obtuvieron imágenes digitales de las hojas seleccionadas con un escáner comercial tipo “flatbed” HP Scanjet G2410, y se almacenaron en el formato JPG con una resolución de 300 dpi. Concluido el escaneo de la hoja, por cada uno de los foliolos, se tomaron seis submuestras con una dimensión de 100 por 100 píxeles. Por cada tratamiento se obtuvieron 90 submuestras, lo que constituye un conjunto de 810 submuestras por cada etapa de desarrollo. El proceso de obtención de las submuestras se llevó a cabo usando el programa de distribución libre para Windows, GIMP 2.8.22 (<https://www.gimp.org>).

Análisis de color y textura

Con el uso del programa en Microsoft® Visual Basic® se obtuvieron los valores numéricos de los píxeles para los tres canales de color del espacio RGB (Rojo, verde y azul, por sus siglas en inglés). Los valores RGB se convirtieron al clasificador de color estándar sRGB (lineales) definido por la Commission Internationale de L'Éclairage (IEC, IEC61966-2-1, 1999 citado por Mendoza *et al.* 2006), con lo que se calculó el espacio de color CIE-Lab. El croma (C) se calculó con la ecuación $(a^2+b^2)^{1/2}$, mientras que el matiz (H) se calculó a partir del arcotangente de la relación a/b (McGuire, 1992), donde a y b son dos canales del espacio de color CIE-Lab. Para el análisis de textura, las subimágenes de 100 por 100 píxeles fueron transformadas a una escala de grises de 8 bits (A). Con base en la metodología de Haralick *et al.* (1973), las imágenes A se cuantizaron a 16 tonos de grises (n). Después se obtuvo la matriz de co-ocurrencia de frecuencias relativas p_{ij} de pares de píxeles vecinos en una imagen digital, uno teniendo nivel de tono i y otro j , separados por $\delta = (r = 1, \theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ)$ en la submuestra, donde r es la distancia en píxeles y θ el ángulo. $\delta = (r, \theta)$ denota un vector en las coordenadas polares de la imagen. Los valores p_{ij} por ángulo se promediaron para generar una matriz de co-ocurrencia promedio, que fue registrada en una hoja de cálculo de n columnas y m hileras de la matriz de tonos de gris de dependencia espacial de la imagen, donde n es el número de tonos de grises (16 en este caso) y $n = m = 16$. Con los p_{ij} se obtuvieron las siguientes características texturales:

Segundo momento angular (SMA):

$$SMA = \sum_{i=1}^n \sum_{j=1}^m [p_{ij}(\delta, \theta)]^2$$

Ec. 1

Entropía (EN):

$$EN = \sum_{i=1}^n \sum_{j=1}^m p_{ij}(\delta, \theta) \log[p_{ij}(\delta, \theta)]$$

Ec. 2

Inercia (IN):

$$IN = \sum_{i=1}^n \sum_{j=1}^m (i-j)^2 p_{ij}(\delta, \theta)$$

Homogeneidad local (HoL):

$$HoL = \sum_{i=1}^n \sum_{j=1}^m 1/[1 + (i-j)^2] p_{ij}(\delta, \theta)$$

Las siguientes características texturales (Ecuaciones 5 a 7) se describen con base en la metodología de Chappard *et al.* (2003). El índice de rugosidad (RI) (Ec. 5) se calculó a partir de las imágenes A . Después se obtuvieron imágenes binarias ($A\sigma$) separadas sucesivamente por umbrales definidos cada 30 niveles de grises (σ). A cada $A\sigma$ se le aplicó una “apertura morfológica” con un elemento estructurante de tres por tres que consiste de aplicar una erosión seguida de una dilación para eliminar píxeles aislados y en seguida se obtuvo su borde en la imagen binaria $B^{A\sigma}$. Se contó el número de píxeles del borde $x(B^{A\sigma})$, y del área Ar a partir de la imagen A para calcula RI (Ec. 5), dónde n es el número de imágenes binarias $A\sigma$.

Índice de rugosidad (IR):

$$IR = \left[\sum_{i=1}^n x(B^{A\sigma})_i \right] / Ar$$

Ec. 5

Heterogeneidad (HE):

$$HE = (A_b + A_w) / (A_b + A_g + A_w)$$

Ec. 6

No-uniformidad del nivel de gris (GLN):

$$GLN = \sum_{i=1}^M \left(\sum_{j=1}^N p(i, j) \right)^2 / s$$

Ec. 7

Respecto a la heterogeneidad (HE) de la imagen (Ec. 6), las imágenes A se mapearon a tres niveles de grises, negro (0 a 85 se mapeó como 0), gris (86 a 171 se mapeó a 150) y blanco (172 a 255 se mapeó a 255). HE se calculó con base en las áreas negra (A_b), gris (A_g) y blanca (A_w) de cada muestra de imagen.

Para calcular la no-uniformidad del nivel de gris (GLN) (Ec. 7) se usaron las imágenes A y se cuantizaron a 16 niveles de grises. Los píxeles consecutivos del mismo valor de gris en una dirección dada constituyen un “camino” (Chu *et al.* 1990). Chappard *et al.* (2003) consideran únicamente dos direcciones, mismas que se utilizan en el presente trabajo: horizontal (GLNH) y vertical (GLNV). En cada dirección se calculó una matriz bidimensional con los elementos p_{ij} que representa el número de caminos con longitud j y de nivel de gris i . Además se calculó s que se refiere al número total de caminos en toda la imagen en una dirección determinada, dónde M es el número de niveles de gris, y N es el camino de mayor longitud en una dirección determinada para un nivel de gris.

Entrenamiento y prueba de escenarios de entradas

El programa en Visual Basic® usó una hoja de cálculo para almacenar los valores promedio por subimagen de los canales de los espacios de color anteriores y de las características texturales (Figura 3.2). Después se usó un archivo de texto delimitado por comas para guardar los datos, con lo cual se obtuvieron 810 datos entrada-salida para cada uno de los cuatro muestreos. También se consideró un quinto caso en el que se juntaron los datos de las cuatro fechas de muestreo (3240 datos). El clasificador utilizado se entrenó y probó con el sistema Neuroshell Classifier® Release 3.0 de AI Trilogy (Ward Systems Group, Inc.) con la estrategia de entrenamiento tipo neuronal supervisada, basado en el algoritmo de correlación en cascada propuesto por Fahlman y Lebiere (1990) y ejemplificado en la Figura 3.1 para uno de los escenarios de entrada propuestos. Este algoritmo inicia con una red neuronal artificial mínima que durante el entrenamiento añade, una por una, nuevas unidades en la capa oculta, lo cual

genera una estructura multicapa. Una vez que se añade a la estructura una nueva unidad en la capa oculta, los pesos del lado de las entradas se hacen constantes por lo cual esta unidad se vuelve un detector de patrones permanente en la red neuronal, y está disponible para producir valores de salida o para crear otros detectores de patrones más complejos. Esta arquitectura se caracteriza por la rapidez para entrenar las redes neuronales artificiales con pocos juegos de datos, y donde un patrón de entrada es clasificado de acuerdo con un número específico de categorías (Ward Systems Group, Inc., 1997-2007). Los escenarios de entrada que se proponen para cada una de las cuatro fechas de muestreo, tuvieron de 9 a 16 variables de color (los canales del espacio de color RGB, los espacios de color CIE-Lab, el C, y el H) y de textura (SMA, EN, IN, HoL, RI, HE, GLNH, y GLNV), nueve clases de salida o tratamientos (Cuadro 3.1), y un número máximo de 150 neuronas en la capa oculta (Figura 3.2). Cuando se evaluaron los cuatro muestreos en conjunto se agregó como una variable de entrada más el número de días después del trasplante (17 variables de entrada). Cada escenario de entrada se repitió 10 veces con particiones aleatorias de los datos de 80 % para el entrenamiento y 20 % para la prueba para todos los casos. La tabla de contingencia del mejor escenario de entrada por fecha de evaluación, se obtuvo con base en la respuesta global del porcentaje de clasificación correcta promedio de las diez particiones aleatorias. También, en el segundo experimento, se presentó la tabla de contingencia del mejor escenario de entrada. Se calculó la sensibilidad, referida como la fracción de observaciones con el síntoma identificado correctamente, y la especificidad, referida como la fracción de observaciones descartadas correctamente de tener el síntoma.

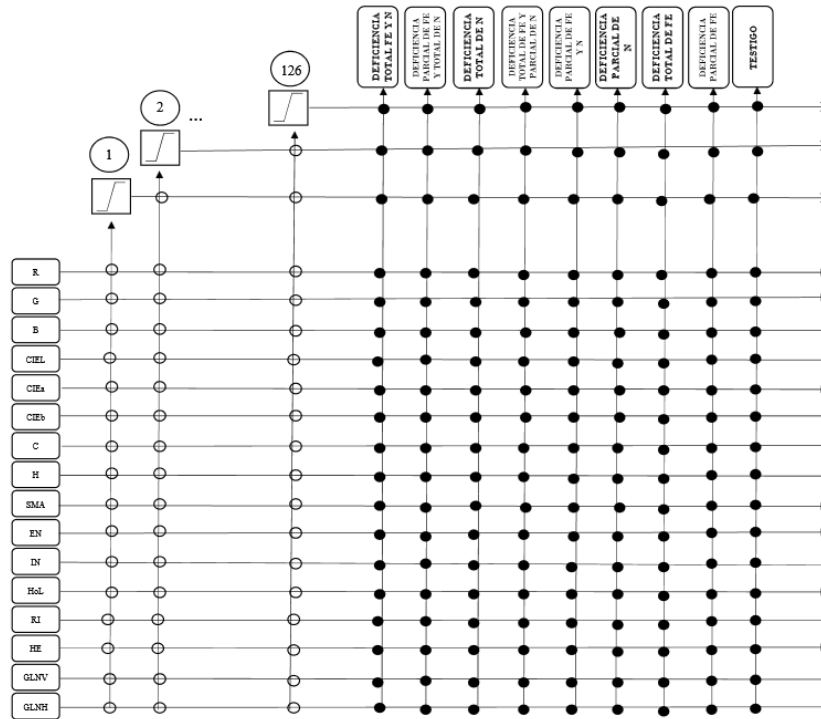


Figura 3.1. Estrategia de entrenamiento tipo neuronal supervisada; esquema basado en el algoritmo de correlación en cascada propuesto por Fahlman y Lebiere (1990) para un clasificador con 16 variables de entrada y nueve variables de salida para identificar interacciones de deficiencias de Fe y N en hojas de frijol var. Negro Chapingo a los 35 ddt. Los números en los círculos indican las neuronas de la capa interna que se agregan una tras otra hasta llegar a 126. Los círculos vacíos indican conexiones congeladas, mientras que los círculos negros son entrenados repetidamente.

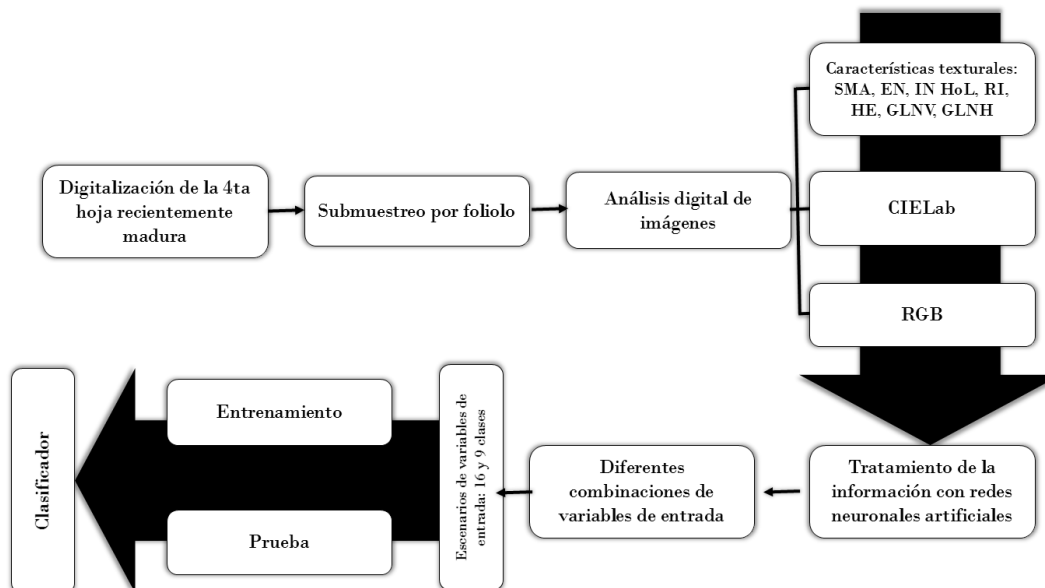


Figura 3.2. Diagrama de procesos abreviados para un clasificador con 16 variables de entrada y nueve variables de salida para identificar interacciones de deficiencias de Fe y N en hojas de frijol var. Negro Chapingo para cada uno de los cuatro muestreos especificados.

Análisis de varianza

El análisis de varianza de los parámetros mencionados en el apartado de Resultados y Discusión, se hicieron con el programa SAS 9.4 para *Windows* (SAS Institute Inc.). La comparación de medias se llevó a cabo con la prueba de Tukey ($\alpha = 0.05$).

3.5 RESULTADOS Y DISCUSIÓN

Los síntomas en hoja de frijol *var. Negro Chapingo* causados por la combinación de niveles de Fe y N en la solución Steiner se identificaron correctamente con porcentaje global de clasificación correcta en la prueba (CCP) de 72 a 81 % (Cuadro 3.2). En general, los CCP más bajos (72.1 – 72.7 %) se obtuvieron en el primer muestreo (35 ddt) y considerando los tres escenarios de entrada propuestos, que incluyeron tanto variables de color como características de textura.

De acuerdo con el análisis de varianza no existen diferencias estadísticas significativas (*valor* $p = <.0001$, $\alpha = 0.005$) entre el uso de alguno de los tres escenarios de entrada propuestos en el Cuadro 3.2. Sin embargo, cuando el análisis de varianza se lleva a cabo tomando en cuenta el número óptimo de neuronas de la capa interna usado para los escenarios de entrada 1, 2 y 3 (126, 104 y 120 respectivamente), se encuentra que sí existen diferencias estadísticas significativas ($p \leq 0.05$). Por lo que, el segundo escenario de entrada (16 variables de entrada) es diferente al resto por el menor número óptimo de neuronas de la capa interna, pero con un

número mayor de variables de entrada en uso, con lo cuál, no necesariamente es menos complejo.

Cuadro 3.2. Entrenamiento y prueba promedio de 10 particiones para crear el clasificador con redes neuronales artificiales y tres escenarios de entrada e identificar nueve clases de salida o niveles de fertilización con base en la textura y color de hojas de frijol: (1) Solución Steiner –N, –Fe; (2) Sol. Steiner –N, -50 % Fe; (3) Sol. Steiner –N; (4) Sol. Steiner –50 % N, –Fe; (5) Sol. Steiner –50 % N, –50 % Fe; (6) Sol. Steiner –50 % N; (7) Sol. Steiner –Fe; (8) Sol. Steiner –50 % Fe y (9) Solución Steiner, Testigo. Muestreo 1 (35 ddt), Muestreo 2 (56 ddt), Muestreo 3 (67 ddt) y Muestreo 4 (78 ddt).

CLSF	MST	Variables de entrada	NN	CCE (%)	CCP (%)	DSTD
1	1	CIEL, CIEa, CIEb, C, H, RI, HE, GLNV, GLNH	126.1	87.9	72.7	3.4
1	2	CIEL, CIEa, CIEb, C, H, RI, HE, GLNV, GLNH	103.6	85.3	73.1	2.7
1	3	CIEL, CIEa, CIEb, C, H, RI, HE, GLNV, GLNH	120.2	91.1	74.5	3.6
1	4	CIEL, CIEa, CIEb, C, H, RI, HE, GLNV, GLNH	117	92.1	76.0	2.1
2	1	R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN HoL, RI, HE, GLNV, GLNH	126.1	87.9	72.0	2.5
2	2	R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN HoL, RI, HE, GLNV, GLNH	103.6	85.3	74.9	2.7
2	3	R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN HoL, RI, HE, GLNV, GLNH	120.2	91.1	77.5	3.1
2	4	R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN HoL, RI, HE, GLNV, GLNH	117	92.1	81.2	3.0
3	1	R, G, B, C, H, RI, HE, GLNV, GLNH	142.6	82.5	72.1	1.4
3	2	R, G, B, C, H, RI, HE, GLNV, GLNH	137	82.8	74.5	3.2
3	3	R, G, B, C, H, RI, HE, GLNV, GLNH	140.3	87.4	75.5	2.5
3	4	R, G, B, C, H, RI, HE, GLNV, GLNH	142.4	86.0	76.0	2.8

CLSF: Clasificador; MST: Muestreo; CCE: Clasificación correcta en el entrenamiento; NN: Número de neuronas de la capa interna; CCP: Clasificación correcta en la prueba; DSTD: Desviación estándar del CCP; R: Rojo, G: Verde, y B: Azul son canales del espacio de color RGB; CIEL: Luminosidad, CIEa y CIEb: Canales de color del espacio de color CIELab; C: croma; H: matiz; SMA: Segundo momento angular; EN: Entropía; IN: Inercia; HoL: Homogeneidad local; RI: Índice de rugosidad; HE: Heterogeneidad; GLNV: No uniformidad del nivel de gris en dirección vertical; GLNH: No uniformidad del nivel de gris en dirección horizontal.

Respecto a los siguientes tres muestreos restantes (56, 67 y 78 ddt), con el análisis de varianza se obtiene que el número de neuronas empleadas por el escenario de entrada 2 es significativo estadísticamente ($p \leq 0.05$) respecto al resto de los escenarios de entrada al usar un número óptimo de neuronas de la capa interna menor.

Por otro lado, cuando se hace una comparación entre el desempeño de los escenarios de entrada y la fecha de muestreo, se tiene que, para el caso del escenario de entrada 1 (CIEL, CIEa, CIEb, C, H, RI, HE, GLNV, GLNH) no mejora el desempeño del mismo a lo largo del tiempo ya que no existen diferencias estadísticas entre las cuatro fechas de muestreo (35, 56, 67 y 78 ddt) a pesar que los síntomas en las hojas podrían ser más evidentes a través del tiempo (Figuras 3.3 y 3.4). No obstante, el CCP es mayor a 72 % en todos los casos.

En el caso del escenario de entrada 3, existen diferencias entre la primera y última fecha de muestreo (35 y 78 ddt), por lo que el desempeño de este escenario de entrada (R, G, B, C, H, RI, HE, GLNV y GLNH) mejora a través del tiempo ($p \leq 0.05$). La diferencia entre el escenario de entrada 3 y el 1 es el uso de dos espacios de color diferente, el RGB y CIELab, respectivamente, mientras que en ambos escenarios se incluyen C, H y las mismas cuatro características texturales. Respecto al escenario de entrada 2, cuando se hizo el análisis de varianza, se encontró que las diferencias estadísticas ($p \leq 0.05$) existen entre las fechas de muestreo 1 (38 ddt), 3 (67), y 4 (78 ddt), lo que significa una mejora en CCP a través del tiempo, y estos cambios se aprecian visualmente al comparar las submuestras por foliolo (Figura 3.3) y la coloración de las hojas (Figura 3.4).

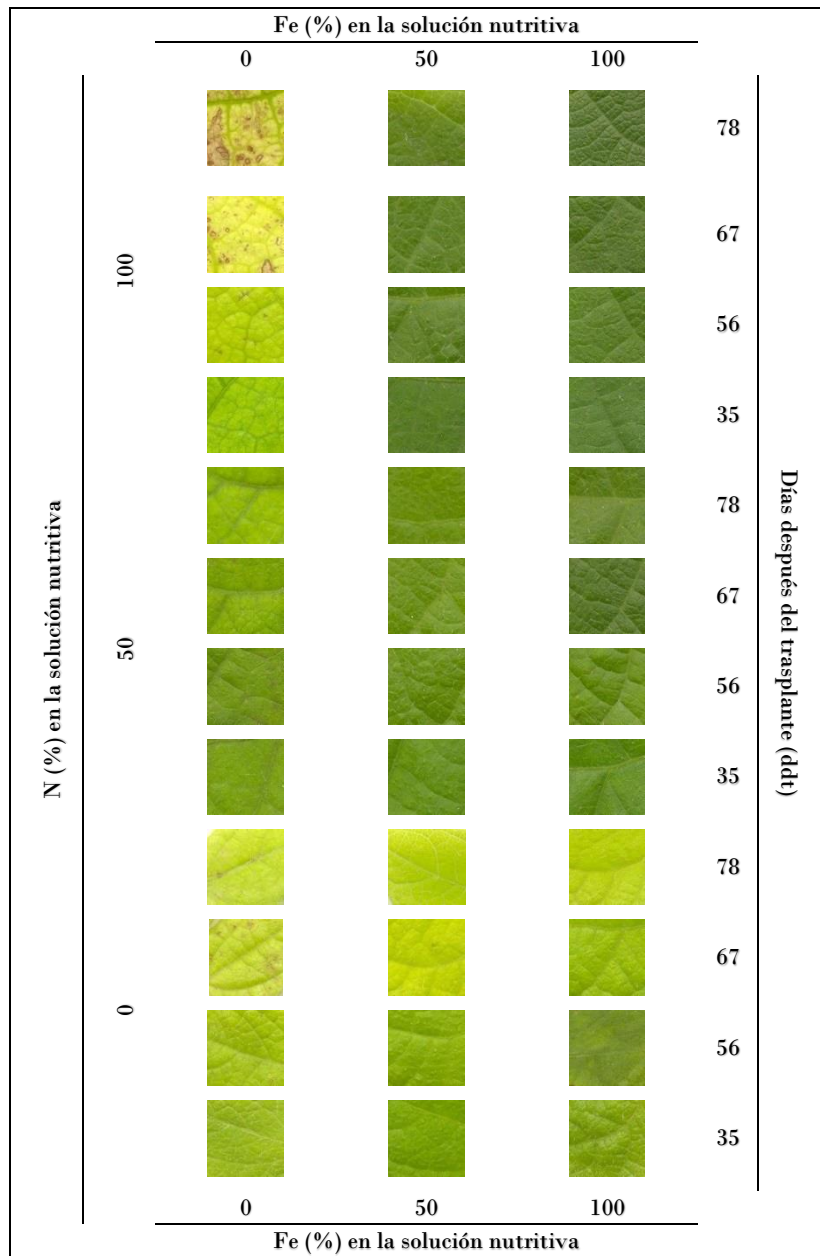


Figura 3.3. Submuestras representativas del foliolo izquierdo de una repetición de cada uno de los tratamientos o clases en los cuatro momentos de muestreo (35, 56, 67 y 78 ddt) de las hojas de frijol *var.* Negro Chapingo.

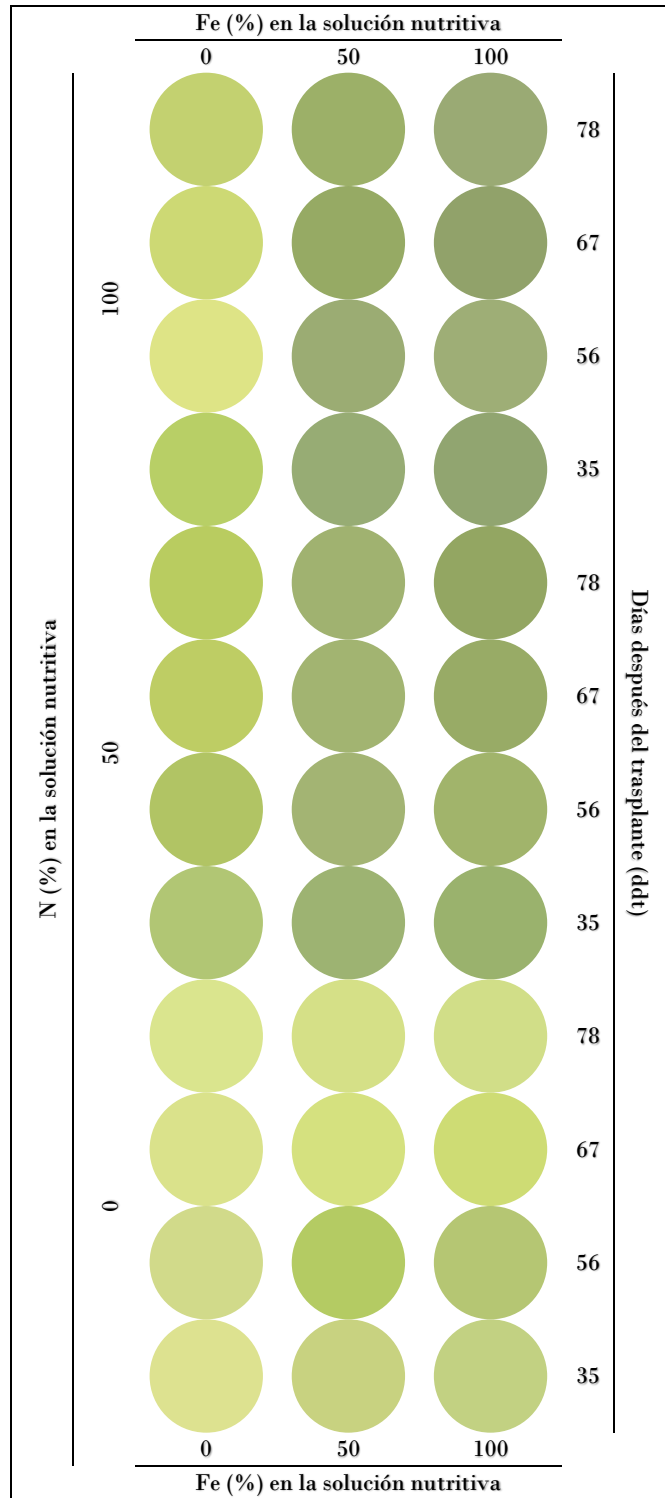


Figura 3.4. Color promedio en el espacio RGB de cada uno de los tratamientos en los cuatro momentos de muestreo (35, 56, 67y 78 ddt) de las hojas de frijol *var.* Negro Chapingo.

Cuadro 3.3. Estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la prueba, y nueve clases de salida: (1) Solución Steiner -N, -Fe; (2) Sol. Steiner -N, -50 % Fe; (3) Sol. Steiner -N; (4) Sol. Steiner -50 % N, -Fe; (5) Sol. Steiner -50 % N, -50 % Fe; (6) Sol. Steiner -50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner, Testigo. Muestreo 1 (35 ddt).

	Observado										Valor predictivo Positivo (%)	
	1	2	3	4	5	6	7	8	9	Total		
Clasificado	1	16.2	0.8	0.4	0	0	0	0.5	0	0.1	18	91
	2	1.1	15.6	2.5	0.4	0	0	0.3	0	0	19.9	79
	3	0.3	0.8	10.5	2.2	0	0.3	0.4	0	0	14.5	77
	4	0	0.6	3.9	14.5	0.1	0.4	1.8	0	0	21.3	69
	5	0	0	0.2	0.2	10.3	7.7	0	1.4	0.1	19.9	52
	6	0	0	0.1	0.7	5.2	8.4	0.1	0.4	0.2	15.1	57
	7	0	0.2	0.1	1.1	0.3	0.7	13.2	0	0	15.6	84
	8	0.1	0.1	0	0	0.7	0.6	0.1	14.4	3.5	19.5	75
	9	0	0	0	0	0.1	0.4	0	4.1	13.6	18.2	74
Total	17.7	18.1	17.7	19.1	16.7	18.5	16.4	20.3	17.5	162		
	0.91	0.87	0.60	0.77	0.62	0.46	0.81	0.71	0.78	Proporción verd.-pos.		
	0.01	0.03	0.03	0.05	0.07	0.05	0.02	0.04	0.03	Proporción falso-pos.		
	0.99	0.97	0.97	0.95	0.93	0.95	0.98	0.96	0.97	Proporción verd.-neg.		
	0.09	0.13	0.40	0.23	0.38	0.54	0.19	0.29	0.22	Proporción falso-neg.		
	91	87	60	77	62	46	81	71	78	Sensibilidad (%)		
	99	97	97	95	93	95	98	96	97	Especificidad (%)		

En los Cuadros 3.3 a 3.6 se presentan las estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la etapa de prueba y con nueve clases de salida. En el primer muestreo, se observa que los porcentajes de clasificaciones correctas se sitúan en un rango de 46 a 91 % en términos de sensibilidad. Las clases con mayor dificultad para ser asignadas son la (5) Sol. Steiner -50 % N, -50 % Fe y (6) Sol. Steiner -50 % N, y la confusión, como puede observarse en el Cuadro 3.3, se da entre estas mismas, 46 y 31 % para 5 y 6 respectivamente. Los patrones de las hojas con tales niveles de deficiencia de Fe o N son pa-

recidos, aunado al factor tiempo en que se hizo el muestreo. Debe considerarse que ambos tienen 50 % menos de nitrógeno en la solución nutritiva, y el efecto del hierro se enmascara y no se observa un patrón en la hoja en consecuencia. La clase (3) Sol. Steiner -N fue otra clase que presentó un porcentaje de clasificación correcta inferior a 70 % y que se confunden con la clase (4) Sol. Steiner -50 % N, -Fe en el orden de 22 % (Cuadro 3.3). Debe considerarse que los síntomas de deficiencia de nitrógeno causados por estos tratamientos aún no se presentaban visualmente en esta fecha de muestreo; además, el muestreo se hizo en las hojas superiores o recientemente maduras, lo cual es lo recomendado para detectar deficiencias de Fe; en cambio, para detectar deficiencias de N se requieren hojas de estratos inferiores. Por otro lado, seis clases se asignaron con un porcentaje de éxito superior a 70 % (Cuadro 3.3). Por lo que puede considerarse que el clasificador obtenido es adecuado para diferenciar las deficiencias simultáneas de Fe y N en hojas de frijol a los 35 ddt. En cuanto a la especificidad, esta es superior a 93 % en todas las clases y fechas de muestreo (Cuadros 3.3 al 3.6).

Después de 21 ddt (fecha de muestreo 2), el CCP aumentó 3%. De manera particular, la sensibilidad de la clase 5 (Sol. Steiner -50 % N, -50 % Fe) aumentó 11 % (Cuadros 3.3 y 3.4). Cuando el nivel de N en la solución nutritiva es del 50 %, la confusión entre las clases (5) y (6) se mantiene entre ambas. Sin embargo, el desempeño del clasificador disminuye 12 % cuando se intenta clasificar el nivel de 50 % de Fe únicamente (clase 8). Esta tendencia se mantiene a lo largo de los muestreos subsecuentes. Existe una mejora en la sensibilidad del 15% en promedio en las clases 1, 3, 4, 5, 6 y 7 (Cuadro 3.4). No obstante, la identificación del tratamien-

to con nutrición óptima (clase 9) se confunden entre sí en el orden del 26 % con la clase 8 (50 % de Fe), por lo que la deficiencia se encuentra en estado latente y sin que se puedan identificar diferencias en los patrones de textura y color.

Cuadro 3.4. Estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la prueba, y nueve clases de salida: (1) Solución Steiner -N, -Fe; (2) Sol. Steiner -N, -50 % Fe; (3) Sol. Steiner -N; (4) Sol. Steiner -50 % N, -Fe; (5) Sol. Steiner -50 % N, -50 % Fe; (6) Sol. Steiner -50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner, Testigo. Muestreo 2 (56 ddt).

	Clase verdadera										Valor predictivo Positivo (%)	
	1	2	3	4	5	6	7	8	9	Total		
Clase predicha	1	17.2	0	0	0	0	0	0.2	0	0	17.4	99
	2	0	14.8	1.2	1.1	0	0.2	0.1	0	0	17.4	85
	3	0	1.4	14.9	4	0	0.4	0	0	0	20.7	72
	4	0	1.3	0.8	10.6	0	0.7	0.1	0	0	13.5	80
	5	0	0	0	0	12.8	5.5	0	2.7	1.1	22.1	58
	6	0	0.4	0.1	2.8	2.8	10.7	0.1	0.4	0.5	17.8	60
	7	0.1	0	0	0	0	0	18.2	0	0	18.3	99
	8	0	0	0	0	0.2	0.8	0	10.8	4.8	16.6	65
	9	0	0	0.1	0	1.7	0.6	0	4.5	11.3	18.2	63
Total	17.3	17.9	17.1	18.5	17.5	18.9	18.7	18.4	17.7	162		
	0.99	0.83	0.88	0.59	0.73	0.57	0.98	0.59	0.65	Proporción verd.-pos.		
	0.00	0.02	0.04	0.02	0.06	0.05	0.00	0.04	0.05	Proporción falso-pos.		
	1.00	0.98	0.96	0.98	0.94	0.95	1.00	0.96	0.95	Proporción verd.-neg.		
	0.01	0.17	0.12	0.41	0.27	0.43	0.02	0.41	0.35	Proporción falso-neg.		
	99	83	88	59	73	57	98	59	65	Sensibilidad (%)		
	100	98	96	98	94	95	100	96	95	Especificidad (%)		

Con una diferencia de 32 y 11 días del muestreo 3 (Cuadro 3.5) con los muestreos 1 y 2 respectivamente, la identificación de una condición normal de nutrición (Clase 9) en el cultivo de frijol puede clasificarse con un porcentaje de éxito de 70 % (Cuadro 3.5). Por lo que a los 67 ddt, podría decirse que la deficiencia parcial de N en las hojas recientemente maduras no se alcanza a percibir de forma consistente debido a que el suministro de este elemento por parte

de las hojas maduras aún intenta mitigar los estragos de una deficiencia parcial de N. Mientras que la identificación de las deficiencias totales de Fe y N mejora debido a que la aparición de los síntomas se intensifica y el efecto de la ausencia de ambos elementos ya es visible.

Cuadro 3.5. Estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la prueba, y nueve clases de salida: (1) Solución Steiner -N, -Fe; (2) Sol. Steiner -N, -50 % Fe; (3) Sol. Steiner -N; (4) Sol. Steiner -50 % N, -Fe; (5) Sol. Steiner -50 % N, -50 % Fe; (6) Sol. Steiner -50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner, Testigo. Muestreo 3 (67 ddt).

		Clase verdadera									Total	Valor predictivo Positivo (%)
		1	2	3	4	5	6	7	8	9		
Clase predicha	1	17.2	0.2	0	0.1	0	0	0	0	0	17.5	98
	2	0	14.9	0.9	0.4	0	0	0.6	0.1	0.1	17	88
	3	0.1	2.4	16	0.7	0	0	0.5	0	0	19.7	82
	4	0	0.4	0.2	15.8	0.2	0.2	0.9	0.2	0	17.9	88
	5	0	0	0	0.3	13.4	4.7	0.1	2.7	0.1	21.3	63
	6	0	0	0	0	1.7	7.6	0.2	3.5	0.3	13.3	58
	7	0	0	0	1.2	0	0	16.1	0	0.1	17.4	92
	8	0	0	0	0	1.5	2.7	0.3	10.2	3	17.7	57
	9	0	0	0	0	0.7	3.7	0	1.7	14.1	20.2	70
Total		17.3	17.9	17.1	18.5	17.5	18.9	18.7	18.4	17.7	162	
		1.00	0.84	0.94	0.87	0.77	0.41	0.87	0.55	0.80		Proporción verd.-pos.
		0.00	0.01	0.03	0.01	0.05	0.04	0.01	0.05	0.04		Proporción falso-pos.
		1.00	0.99	0.97	0.99	0.95	0.96	0.99	0.95	0.96		Proporción verd.-neg.
		0.01	0.16	0.06	0.13	0.23	0.59	0.13	0.45	0.20		Proporción falso-neg.
		100	84	94	87	77	41	87	55	80		Sensibilidad (%)
		100	99	97	99	95	96	99	95	96		Especificidad (%)

En términos de sensibilidad, para la clase 5 (deficiencia parcial de Fe y N), el porcentaje es 77 %, el cual está asociado con la proporción verdadero-positiva del clasificador y que corresponde a aquellas muestras dentro del 20 % que realmente pertenecen a esa clase y que se asignaron correctamente en la etapa de prueba. De manera específica, para esta clase, los

porcentajes de sensibilidad aumentaron del muestreo 1, 2 y 3 (62, 73 y 77 % respectivamente), sin embargo, tal porcentaje decrece (63 %) en la última fecha de muestreo (Cuadro 3.6).

En el último de los muestreos, con 11 días más respecto al tercer muestreo, la tendencia en los valores CCP se mantienen. Aunque este porcentaje aumenta casi 10 % respecto al primer muestreo (Cuadros 3.2 y 3.6), la identificación de niveles parciales de Fe o N en la solución nutritiva genera valores de éxito moderados (72 y 63 % respectivamente).

Cuadro 3.6. Estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH en la prueba, y nueve clases de salida: (1) Solución Steiner -N, -Fe; (2) Sol. Steiner -N, -50 % Fe; (3) Sol. Steiner -N; (4) Sol. Steiner -50 % N, -Fe; (5) Sol. Steiner -50 % N, -50 % Fe; (6) Sol. Steiner -50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner, Testigo. Muestreo 4 (78 ddt).

		Clase verdadera									Valor predictivo	
		1	2	3	4	5	6	7	8	9	Total	Positivo (%)
Clase predicha	1	17.3	0.2	0	0	0	0	0.1	0	0	17.6	98
	2	0	15.9	0.8	0.9	0	0	0.5	0	0	18.1	88
	3	0	0.9	16.5	0.7	0.2	0	0.4	0	0	18.7	87
	4	0	0.6	0.2	15.9	0.1	0	0.4	0	0	17.2	93
	5	0	0	0	0.2	11	1.7	0	1.4	1	15.3	72
	6	0	0	0.1	0	2.2	11.6	0.3	3.4	1.5	19.1	62
	7	0	0.1	0	0.3	0.1	0.6	17.1	0	0	18.2	95
	8	0	0	0.1	0.1	1.9	3.4	0	12.9	1.2	19.6	66
	9	0	0	0	0.1	1.9	1.6	0	0.6	14	18.2	78
Total		17.3	17.7	17.7	18.2	17.4	18.9	18.8	18.3	17.7	162	
		1.00	0.89	0.92	0.88	0.63	0.63	0.91	0.72	0.79	Proporción verd.-pos.	
		0.00	0.02	0.02	0.01	0.03	0.05	0.01	0.05	0.03	Proporción falso-pos.	
		1.00	0.98	0.98	0.99	0.97	0.95	0.99	0.95	0.97	Proporción verd.-neg.	
		0.00	0.11	0.08	0.12	0.37	0.37	0.09	0.28	0.21	Proporción falso-neg.	
		100	89	92	88	63	63	91	72	79	Sensibilidad (%)	
		100	98	98	99	97	95	99	95	97	Especificidad (%)	

Por otro lado, los porcentajes de sensibilidad para todos los tratamientos y considerando la última fecha de muestreo, son superiores al 60 % e incluyen a la deficiencia total de N y su interacción con los tres niveles de Fe: 0, 50 y 100 % (Cuadro 3.6). Es decir, la ausencia de N tiene un efecto definido en la determinación de los patrones tanto de textura como de color en la primera hoja recientemente madura de frijol. La sensibilidad de la clase 9 (testigo) aumenta 63 a 78 % conforme el contraste entre los tratamientos es más evidente a través del tiempo (Cuadros 3.3 y 3.6). En los escenarios de entradas (Cuadro 3.7), cuando se incluye la fecha de muestreo como una variable adicional a los tres escenarios de entrada propuestos anteriormente (Cuadro 3.2), se tiene que el desempeño evaluado como CCP es inferior. En este caso, el escenario de entrada 2 (Cuadro 3.7) tiene una menor desviación estándar que cuando se obtuvo por fecha de muestreo.

Cuadro 3.7. Entrenamiento y prueba promedio de 10 particiones para crear el clasificador con redes neuronales artificiales y tres escenarios de entrada e identificar nueve clases de salida o niveles de fertilización con base en la textura y color de hojas de frijol: (1) Solución Steiner -N, -Fe; (2) Sol. Steiner -N, -50 % Fe; (3) Sol. Steiner -N; (4) Sol. Steiner -50 % N, -Fe; (5) Sol. Steiner -50 % N, -50 % Fe; (6) Sol. Steiner -50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner, Testigo. Muestreo 1 (35 ddt), Muestreo 2 (56 ddt), Muestreo 3 (67 ddt) y Muestreo 4 (78 ddt).

CLSF	MST	Variables de entrada	NN	CCE (%)	CCP (%)	DSTD
1	Todos	CIEL, CIEa, CIEb, C, H, RI, HE, GLNV, GLNH, DDT	144.0	68.5	63.1	1.9
2	Todos	R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN HoL, RI, HE, GLNV, GLNH, DDT	145.6	73.5	65.7	2.0
3	Todos	R, G, B, C, H, RI, HE, GLNV, GLNH, DDT	147.7	68.8	63.2	2.6

CLSF: Clasificador; MST: Muestreo; CCE: Clasificación correcta en el entrenamiento; NN: Número de neuronas de la capa interna; CCP: Clasificación correcta en la prueba; DSTD: Desviación estándar del CCP; R: Rojo, G: Verde, y B: Azul son canales del espacio de color RGB; CIEL: Luminosidad, CIEa y CIEb: Canales de color del espacio de color CIELab; C: croma; H: matiz; SMA: Segundo momento angular; EN: Entropía; IN: Inercia; HoL: Homogeneidad local; RI: Índice de rugosidad; HE: Heterogeneidad; GLNV: No uniformidad del nivel de gris en dirección vertical; GLNH: No uniformidad del nivel de gris en dirección horizontal; DDT: Días después del transplante.

Cuadro 3.8. Estadísticas promedio (n= 10) del arreglo matricial o tabla de contingencia para el escenario de entrada: R, G, B, CIEL, CIEa, CIEb, C, H, SMA, EN, IN, HoL, RI, HE, GLNV, GLNH y DDT en la prueba, y nueve clases de salida: (1) Solución Steiner -N, -Fe; (2) Sol. Steiner -N, -50 % Fe; (3) Sol. Steiner -N; (4) Sol. Steiner -50 % N, -Fe; (5) Sol. Steiner -50 % N, -50 % Fe; (6) Sol. Steiner -50 % N; (7) Sol. Steiner -Fe; (8) Sol. Steiner -50 % Fe y (9) Solución Steiner, Testigo. Datos en conjunto de prueba y entrenamiento del Muestreo 1 (35 ddt), Muestreo 2 (56 ddt), Muestreo 3 (67 ddt) y Muestreo 4 (78 ddt).

		Clase verdadera									Total	Valor predictivo Positivo (%)
		1	2	3	4	5	6	7	8	9		
Clase predicha	1	66.1	6.8	1.1	0.6	0	0	2.6	0.2	0	77.4	85
	2	1.6	45.1	9	4	0.2	0.4	1.3	0	0	61.6	73
	3	1.2	13.5	49	7.5	0.5	1.1	1.3	0.4	0	74.5	66
	4	0	5.5	14.4	54.4	1.8	4.2	7.1	0.9	0	88.3	62
	5	0	0.2	0.7	1	44.6	24.6	0.8	9.4	3.4	84.7	53
	6	0.1	0	0.7	1.5	11.8	26.3	0.2	11.3	6.8	58.7	45
	7	0.9	0.6	1.7	2.2	1.2	0.3	56.9	0	0	63.8	89
	8	0	0	0.2	0.1	4.8	12.7	0.1	33.2	12.1	63.2	53
	9	0	0	0	0.1	7	5.5	0.2	12.7	50.3	75.8	67
Total		69.9	71.7	76.8	71.4	71.9	75.1	70.5	68.1	72.6	648	
		0.95	0.63	0.64	0.76	0.62	0.35	0.81	0.49	0.70		Proporción verd.-pos.
		0.02	0.03	0.04	0.06	0.07	0.06	0.01	0.05	0.04		Proporción falso-pos.
		0.98	0.97	0.96	0.94	0.93	0.94	0.99	0.95	0.96		Proporción verd.-neg.
		0.05	0.37	0.36	0.24	0.38	0.65	0.19	0.51	0.30		Proporción falso-neg.
		95	63	64	76	62	35	81	49	70		Sensibilidad (%)
		98	97	96	94	93	94	99	95	96		Especificidad (%)

En el Cuadro 3.8, se presentan las estadísticas promedio (n=10) del arreglo matricial en la prueba del escenario de entrada 2 y agregando como variable de entrada adicional el tiempo de muestreo. El valor predictivo positivo es superior a 62 % para seis clases de nueve, siendo el valor superior 85 %. Por otro lado, la clase seis con mayor problema para asignarse a un tratamiento y se confunde con la clase cinco. Ambas tienen en común que hay una deficiencia de N (50 %), con la diferencia que la clase cinco, esta deficiencia además está en interacción con una deficiencia parcial de Fe (50 %). Se refuerza lo dicho anteriormente, en la que el pa-

trón de textura y color de una deficiencia de N dificulta al clasificador la diferenciación de ambas condiciones nutrimentales, aunque ambas señalen una condición de deficiencia parcial de N.

Por otro lado, la deficiencia parcial de Fe (-50 % Fe) tiende a confundirse con el testigo (Clase 9) en un porcentaje del 18 %. Probablemente por la condición de latencia de la deficiencia en cuestión y la ya mencionada confusión con la clase (5). En ensayos no presentados se unieron las clases que mayor confusión presentan (clases 5 y 6 por un lado, y clases 8 y 9, por otro). El CCP del escenario de entradas 2 (Cuadro 3.7) aumentó a más de 90 %, con lo que se demuestra la similitud entre los patrones de textura y color entre estas clases.

Cuando se consideran grupos de porcentajes de clasificaciones correctas, los tres niveles de Fe (Clases 7, 8 y 9, Cuadro 3.8) sin interacción alguna con las deficiencias de N, pueden clasificarse correctamente en el orden del 67 % contrastado con el 56 % con el que se clasificarían los tres niveles de N sin interacción alguna con el Fe (Clases 3, 6 y 9, Cuadro 3.8) y cuando se consideran los patrones de textura y color de hojas recientemente maduras. Sin embargo, si se toman en cuenta las fechas de muestreo de manera separada: 35, 56, 67 y 78 ddt, los porcentajes son superiores en todos los casos: 77, 74, 74 y 80 % de clasificaciones correctas cuando se clasifican los tres niveles de Fe sin interacción con niveles de N. Por otro lado, los tres niveles de N (Clases 3, 6 y 9, Cuadros 3.3 a 3.6), se clasifican exitosamente en el orden de 61, 70, 72 y 78 % a los 35, 56, 67 y 78 ddt.

Se sugiere el empleo de otras características morfológicas de la planta para robustecer el clasificador, tales como el área de la hoja. Adicionalmente, evaluar otras características de textura o color que mejoren el desempeño del clasificador para identificar los patrones de las deficiencias parciales tanto de Fe como N y tomando en cuenta el factor tiempo.

3.6 CONCLUSIONES

Se obtuvo un clasificador capaz de identificar niveles de Fe y N en hojas de frijol en cuatro puntos del desarrollo del cultivo con CCP de hasta 81 % en función del momento del muestreo. El desempeño del clasificador se ve afectado, dado que existe un efecto de interacción del nivel de N en el cultivo de frijol en la detección de deficiencias de Fe, de manera importante cuando los niveles de deficiencia son, en proporción, la mitad del requerimiento óptimo nutrimental de estos nutrimentos.

La conclusión principal de este estudio entonces, es que los clasificadores propuestos fueron capaces de diferenciar una exactitud de 63 a 66 % las deficiencias de Fe y/o N y sus interacciones cuando se consideraron las cuatro fechas de muestreo en conjunto, y hasta 76 % cuando se consideran los muestreos de manera separada. La sensibilidad de para el primer caso fue 66 % y 77 % en promedio para fechas de muestreo separadas. La especificidad del clasificador, por otro lado, fue muy alta para ambos casos 96 y 97 %.

3.7 LITERATURA CITADA

- Abadía, J., M. Tagliavini, R. Grasa, R. Belkhodja, A. Abadía, and M. Sanz. 2000. Using the flower Fe concentration for estimating chlorosis status in fruit tree orchards: a summary report. *Journal of Plant Nutrition* 23:2023-2033.
- Chen, L., L. Lin, G. Cai, Y. Sun, T. Huang, K. Wang, and J. Deng. 2014. Identification of nitrogen, phosphorus, and potassium deficiencies in rice based on static scanning technology and hierarchical identification method. *PLoS ONE* 9(11). <https://doi.org/10.1371/journal.pone.0113200>
- Ciganda, V., A. Gitelson, and James Schepers. 2009. Non-destructive determination of maize leaf and canopy chlorophyll content. *Journal of Plant Physiology* 166 (2): 157-167.
- Ercoli, L., M. Mariottia, A. Masoni, and F. Massantinia. 1993. Relationship between nitrogen and chlorophyll content and spectral properties in maize leaves. *European Journal of Agronomy* 22:(2) 113-117.
- Gitelson, A.A., Gritz, Y., and Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology* 160:271–282.
- Kalacska, M., M. Lalonde, and T. R. Moore. 2015. Estimation of foliar chlorophyll and nitrogen content in an ombrotrophic bog from hyperspectral data: Scaling from leaf to image. *Remote Sensing of Environment* 169: 270-279.
- Kergoat, L., Lafont, S., Arneth, A., Le Dantec, V., and Saugier, B., 2008. Nitrogen controls plant canopy light-use-efficiency in temperate and boreal ecosystems. *Journal of Geophysical Research* 113: 1–19.
- Köseoğlu, A. T. and V. Açıkgoz. 1995. Determination of iron chlorosis with extractable iron analysis in peach leaves. *Journal of Plant Nutrition* 18(1):153-161.

- Mendoza, F., P. Dejmek, and L. Aguilera. 2006. Calibrated color measurements of agricultural foods using image analysis, *Postharvest Biology and Technology* 41 (3): 285-295.
- Razeto, B. y G. Valdéz. 2006. Análisis de hierro soluble en tejidos para diagnosticar el déficit de hierro en nectarino. *Agricultura Técnica* 66(2):216-220.
- Schlemmer, M., A. A. Gitelson, J. S. Schepers, R. B. Ferguson, and Y. Peng. 2013. Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation* 25:47-54.
- Story, D., M. Kacira, C. Kubota, A. Akoglu, L. An. 2010. Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments. *Computers and Electronics in Agriculture* 74: 238-243.

CONCLUSIONES GENERALES



La predicción en la agricultura ha sido ampliamente estudiada con la finalidad de anticiparse a los fenómenos que intervienen en la obtención del potencial óptimo del cultivo; en el ámbito de la nutrición, anticiparse a la solución de los desórdenes nutrimentales tiene una relevancia fundamental. Precisamente en esta investigación se plantearon tres hipótesis desde ese enfoque, para la identificación de condiciones nutrimentales de hierro (Fe), manganeso (Mn) y nitrógeno (N); tales hipótesis se confirmaron dados los resultados que se obtuvieron.

En un primer caso, se concluye que existe una correlación de Pearson de 0.71 entre la estimación del contenido de clorofila hecha por el SPAD 502 (Minolta Corp.) y su distribución en el tiempo con el estatus nutrimental de hierro en el cultivo de frijol y haciendo uso de redes neuronales bayesianas regularizadas como un modelo cuantitativo. Mientras que cuando se hace uso de árboles de decisión, como un modelo cualitativo, el porcentaje global de clasificaciones correctas es del 44 %. Por lo que el uso de las variables de entrada propuestas es insuficiente para el empleo de este tipo de modelo, no así para el modelo planteado con redes neuronales bayesianas regularizadas.

También, se demostró que es posible identificar deficiencias de Fe y Mn de frijol con una eficiencia de 70 80 % (porcentaje global de clasificación correcta en la prueba) a partir de imágenes de hojas, mediante un clasificador entrenado con redes neuronales artificiales y que usó

las características texturales y de color de las imágenes como variables de entrada para generar patrones de reconocimiento de las deficiencias en una etapa temprana cuando visualmente no eran perceptibles. La variación en la eficiencia se debe a la diversidad de posibles patrones generados por la red neuronal al cambiar el número de variables de entrada de textura y color y el número de neuronas en la capa interna.

Las deficiencias tanto de Fe como de Mn se observan generalmente en las hojas superiores de cualquier planta por lo que en esta tesis se partió del supuesto de que su interacción puede generar patrones particulares de textura y color dependiendo de la concentración y disponibilidad de cada uno de estos nutrientes. Ambos elementos provocan una clorosis en la hoja. El Fe en particular, es elemento constituyente de muchas enzimas y participa en varias rutas metabólicas por lo que su deficiencia en plantas se demuestra como una clorosis intervenal. Ahora, se desconoce de qué manera se expresa este patrón de clorosis en la hoja superior de la planta de frijol cuando se presenta simultáneamente un nivel de deficiencia de N, elemento básico para la formación de clorofila y que generalmente se identifica como clorosis en las hojas inferiores. En el tercer experimento, se concluyó que las deficiencias de Fe y N generaron patrones específicos de color y textura en las hojas superiores de frijol con porcentaje global de clasificación correcta en la prueba de 72 %, en una etapa temprana cuando los síntomas no son aparentes, hasta 81 % en una etapa avanzada de los síntomas.

Con un desempeño moderado, se puede afirmar que es posible la identificación de condiciones nutrimentales de Fe, Mn o N en el cultivo de frijol en una etapa temprana, cuando la apari-

ción de los síntomas no es visible; se sugiere la evaluación del crecimiento y rendimiento del cultivo para cuantificar el efecto negativo de las deficiencias o excesos en tiempo de muestreo y al mismo tiempo, conocer el grado en que el daño ocasionado por las deficiencias o excesos nutrimentales es reversible.