

TABLA DE CONTENIDOS

	página
Dedicatoria	I
Agradecimientos	II
Tabla de Contenidos	III
Índice de Figuras	VI
Índice de Tablas	X
Resumen	XI
Abstract	XII
1. Introducción	1
1.1. Descripción del Contexto	2
1.2. Objetivo general	3
1.3. Objetivos específicos	3
1.4. Alcances	4
2. Marco teórico	5
2.1. Imágenes hiperespectrales	5
2.2. Cámara hiperespectral y medidor de penetración	6
2.3. Análisis de texturas en trabajos realizados	8
2.3.1. Imágenes hiperespectrales para la textura de la carne	8
2.3.2. Imágenes hiperespectrales para determinación no destructiva de algunos atributos de calidad de frutilla.	9
2.4. Análisis de textura	9
2.5. Construcción de la GLCM	10
2.6. Descriptores de texturas	13
2.7. Visualización de datos de alta dimensionalidad	15
2.8. Interpolación	18
2.9. Errores	19

2.9.1. Error medio	19
2.9.2. Error medio absoluto	20
2.9.3. Raíz del error medio cuadrático	20
2.10. Validación cruzada (o cross-validation)	20
2.10.1. Validación cruzada “Dejando uno fuera” (“leave-one-out cross-validation” o LOOCV)	22
3. Metodología	23
3.1. Archivos BIL, BSQ, BIP	23
3.2. Hypercube	25
3.3. Herramientas	26
3.4. Espectro de luz	27
3.5. División de píxeles	28
3.6. Código	30
3.7. Predicción de firmeza	31
3.8. Clasificación de madurez	33
4. Resultados	35
4.1. GLCM	35
4.2. Visualización de madurez	39
4.3. Predicción de firmeza (errores)	41
5. Conclusiones	44
5.1. Trabajos futuros	45
Glosario	46
Bibliografía	47
Anexos	
A: Análisis de alta dimensionalidad	50
A.1. Imágenes de la variedad camarosa	50
A.2. Imágenes de la variedad camino real	53
A.3. Imágenes de la variedad cruzamiento ₀₃₈₈	57

B: Datos de las variedades: clasificación, peso, firmeza	61
---	-----------

ÍNDICE DE FIGURAS

	página
2.1. Interpretación de las imágenes hiperespectrales, el hipercubo con n bandas y con dos espacios X, Y y una dimensión de longitud de onda.	6
2.2. Componentes del sistema que captura las imágenes hiperespectrales, estos son la cámara, el espectrógrafo, iluminación con fuente led y un computador.	7
2.3. Medidor para ejercer presión llamado penetrómetro, existen presiones distintas para cada instrumento (Kg/cm^2) o (<i>gramos/Newton</i>).	8
2.4. Direcciones que se realiza el análisis GLCM, con distancia D de 1, y Ángulos $0^\circ, 45^\circ, 90^\circ, 135^\circ$	10
2.5. Imagen de prueba de dimensión 4×4 con 4 valores de píxeles.	10
2.6. Todas las posibles combinaciones de los 4 niveles de gris de la imagen de prueba.	11
2.7. Matriz GLCM obtenida en dirección 0° .	11
2.8. Matriz simétrica, obtenida con la suma de la matriz GLCM y su transpuesta.	12
2.9. Matriz normalizada obtenida mediante la ecuación de probabilidad.	13
2.10. Una imagen de muestra con sus distintos descriptores de texturas.	14
2.11. En las dimensiones 1, 2 y 3 es fácil analizar datos, en una dimensión mayor es muy complicado.	15
2.12. Interpretación de extracción de característica, donde $D1, D2, D3, D4$ y $D5$ son las entradas, $PC1$ y $PC2$ las salidas.	16
2.13. Interpretación de la función de selección, donde $D1, D2, D3, D4$ y $D5$ son las entradas, $D1$ y 42 las salidas.	17
2.14. Ejemplo de una interpolación línea. Recta $y = r(x)$ que calcula los valores intermedios en lugar de la función $y = f(x)$	19
2.15. Interpretación de validación cruzada simple (o cross-validation), un conjunto se divide en dos subconjuntos, llamados entrenamiento y prueba.	21
2.16. Interpretación de validación cruzada k -veces (o cross-validation k -fold), un conjunto se divide k veces, dependiendo del tamaño de la muestra.	21

2.17. LOOCV, el conjunto se divide $n - 1$ que es para entrenamiento y 1 sola muestra para prueba. se realiza n veces el procedimiento.	22
3.1. Bandas 1, 2 y 3 para ejemplificar los formatos.	23
3.2. Interpretación del formato BIL.	24
3.3. Interpretación del formato BQS.	24
3.4. Interpretación del formato BIP.	25
3.5. Visualización del archivo BIL en el software HyerCube.	26
3.6. Lectura del archivo para obtener los píxeles para la construcción de la matriz GLCM.	29
3.7. Matrices obtenidas con la división de píxeles con valore 10,50,100,500.	29
3.8. Modelos que servirán para predecir firmeza. Uno basado en el peso y el otro en los descriptores de texturas con los 6 análisis distintos. . . .	31
3.9. Errores obtenidos mediante el proceso Gaussiano, realizado por cada validación cruzada (Dejando uno fuera). Los que se guardan en un arreglo para el término de las iteraciones se entregan los errores MA, MAE Y RMSE.	32
3.10. Visualización de la clasificación de madurez. SG: significa que la fruta aún es verde, LG: significa que la fruta creció pero aún es verde, T: significa que la fruta esta a punto de cambiar a un color rojo, R75: significa que la frutilla está madura al 75 %, R100: significa que la frutilla está madura al 100 %.	34
4.1. Obtención de los descriptores de texturas para cada ángulo en cada capa.	36
4.2. Imagen de escala de gris de una capa de la variedad camino real. . . .	37
4.3. Imagen de escala de gris de una capa de la variedad cruzamiento ₃₈₈ . . .	37
4.4. Línea de píxeles sacada en cada imagen para comparar la reflectancia.	38
4.5. Comparación de reflectancia de las imágenes. En el eje Y es la reflectancia, en el eje X número de píxel de la línea sacada.	38
4.6. Visualización de agrupamientos realizado sin división en los píxeles, SG: significa que la fruta aún es verde, LG: significa que la fruta creció pero aún es verde, T: significa que la fruta esta a punto de cambiar a un color rojo, R75: significa que la frutilla está madura al 75 %, R100: significa que la frutilla está madura al 100 %.	39

4.7. Visualización de agrupamientos realizado con división de 10 en cada píxel, SG: significa que la fruta aún es verde, LG: significa que la fruta creció pero aún es verde, T: significa que la fruta esta a punto de cambiar a un color rojo, R75: significa que la frutilla está madura al 75 %, R100: significa que la frutilla está madura al 100 %.	40
A.1. Análisis realizado a la variedad camarosa, realizando la división de 50 en cada píxel de cada capa.	51
A.2. Análisis realizado a la variedad camarosa, realizando la división de 100 en cada píxel de cada capa.	51
A.3. Análisis realizado a la variedad camarosa, realizando la división de 500 en cada píxel de cada capa.	52
A.4. Análisis realizado a la variedad camarosa, realizando la división de 1000 en cada píxel de cada capa.	52
A.5. Análisis realizado a la variedad camino real, sin realizar división en los píxeles.	53
A.6. Análisis realizado a la variedad camino real, realizando la división de 10 en cada píxel de cada capa.	54
A.7. Análisis realizado a la variedad camino real, realizando la división de 50 en cada píxel de cada capa.	54
A.8. Análisis realizado a la variedad camino real, realizando la división de 100 en cada píxel de cada capa.	55
A.9. Análisis realizado a la variedad camino real, realizando la división de 500 en cada píxel de cada capa.	55
A.10. Análisis realizado a la variedad camino real, realizando la división de 1000 en cada píxel de cada capa.	56
A.11. Análisis realizado a la variedad cruzamiento ₃₈₈ , sin realizar división en los píxeles.	57
A.12. Análisis realizado a la variedad cruzamiento ₃₈₈ , realizando la división de 10 en cada píxel de cada capa.	58
A.13. Análisis realizado a la variedad cruzamiento ₃₈₈ , realizando la división de 50 en cada píxel de cada capa.	58
A.14. Análisis realizado a la variedad cruzamiento ₃₈₈ , realizando la división de 100 en cada píxel de cada capa.	59

A.15. Análisis realizado a la variedad cruzamiento ₃₈₈ , realizando la división de 500 en cada píxel de cada capa.	59
A.16. Análisis realizado a la variedad cruzamiento ₃₈₈ , realizando la división de 1000 en cada píxel de cada capa.	60

ÍNDICE DE TABLAS

	página
3.1. Tiempos de ejecución en niveles de grises, con la división de píxeles y sin división de píxeles.	30
4.1. Los errores MAE y RMSE para la división de 1000, se esperaba un valor más alto referente al sin división de píxel en la predicción de firmeza.	41
4.2. Los errores MAE y RMSE para la división de 1000, se esperaba un valor más alto referente al sin división de píxel, y el modelo sin división que fuera mucho más bajo al modelo basado en peso.	42
4.3. Los errores MAE y RMSE para el modelo sin división son más alto al modelo basado en el peso, y además el modelo de división de 50 es más bajo al modelo sin división.	42
B.1. Datos variedad camarosa: clasificación, peso, firmeza	62
B.2. Datos variedad camino real: clasificación, peso, firmeza	63
B.3. Datos variedad cruzamiento ₃₈₈ : clasificación, peso, firmeza	64