

SISTEMA DE RECONOCIMIENTO DE BILLETES PARA PERSONAS CON DISCAPACIDAD VISUAL MEDIANTE VISIÓN ARTIFICIAL

KELLY JOHANNA TAMAYO ZAPATA

Trabajo de grado para optar al título de Ingeniera Mecatrónica

**Gustavo Andrés Moreno Hincapié, Docente en
Automatización y Robótica**



**UNIVERSIDAD EIA
INGENIERÍA MECATRÓNICA
ENVIGADO
2018**

A todos los niños como Angel, de ojos sin luz y con destellos en el alma.

AGRADECIMIENTOS

A Dios, quien siempre ha estado a mi lado.

A Santi, familia y amigos por su resistencia y apoyo incondicional ante mis retos.

A Geoffrey Acevedo, Director de Mecatrónica, por su diligencia y cooperación.

A mi profesor Gustavo Andrés Moreno Hincapié, por incentivar mi pasión por la Visión Artificial

A todos los Directivos, profesores y compañeros que intervinieron en mi crecimiento académico y personal.

CONTENIDO

	pág.
INTRODUCCIÓN.....	13
1. PRELIMINARES.....	17
1.1 Planteamiento del problema	17
1.2 Objetivos del proyecto	19
1.2.1 Objetivo General.....	19
1.2.2 Objetivos Específicos	19
1.3 Marco de referencia.....	20
1.3.1 Marco Contextual	20
1.3.2 Marco Teórico	23
2. METODOLOGIA.....	54
2.1 Diseño conceptual	54
2.1.1 Entrevista a fundaciones	54
2.2 Especificaciones Objetivo	54
2.2.1 Lista de necesidades	54
2.2.2 Lista de Medidas iniciales.....	55
2.2.3 Lista de medidas de la competencia.....	55
2.2.4 Lista de medidas ideales y aceptables	57
2.2.5 Matriz de necesidades y medidas.....	57
2.3 Diseño conceptual	58
2.3.1 Aclaración del problema	58
2.3.2 Registro de soluciones	59
2.3.3 Diseño de detalle.....	61
3. PRESENTACIÓN Y DISCUSIÓN DE RESULTADOS.....	83

3.1	Prototipo funcional.....	83
3.1.1	Construcción prototipo 1.....	83
3.1.2	Construcción prototipo 2.....	93
3.2	Discusión de metodología	¡Error! Marcador no definido.
4.	CONCLUSIONES.....	¡ERROR! MARCADOR NO DEFINIDO.
5.	REFERENCIAS	¡ERROR! MARCADOR NO DEFINIDO.

LISTA DE TABLAS

Tabla 1. Billetes colombianos en circulación a partir de 2016 (Banco de la República - Colombia, s.f.)	22
Tabla 2. Billetes colombianos en circulación para ser retirados (Banco de la República - Colombia, s.f.)	22
Tabla 3. Lista de necesidades	54
Tabla 4. Lista de medidas.....	55
Tabla 5. Lista de medidas de la competencia	55
Tabla 6. Lista de medidas ideales y aceptables.....	57
Tabla 7. Matriz de necesidades/ medidas.....	57
Tabla 8. Definición de conceptos	60
Tabla 9. Selección de concepto	60
Tabla 10. Características de fuente de alimentación.....	61
Tabla 11. Características de la cámara.....	62
Tabla 12. Características sistema embebido	63
Tabla 13. Distribución de clases y características.....	76
Tabla 14. Matriz de confusión.....	77
Tabla 15. Ejemplo de matriz de confusión para algunos billetes.....	78
Tabla 16. Análisis de la matriz de confusión anterior para una clase	78
Tabla 17. Clase con su respectivo billete.....	90
Tabla 18. Clase con su respectiva etiqueta.....	91
Tabla 19. Porcentaje de exactitud por tipo de característica	92
Tabla 20. Rendimiento de la red.....	106

LISTA DE FIGURAS

	pág.
Ilustración 1. Diferencia de braille entre billete de 10.000 nuevo y deteriorado (Banco de la República- Colombia, 2013)	18
Ilustración 2. Ojo humano (Clínica RAHHAL, s.f.)	23
Ilustración 3. Subsistema de las imágenes (Universidad Politécnica de Madrid).....	24
Ilustración 4. Esquema de las relaciones entre visión por computador y relación (Academic, s.f)	24
Ilustración 5. Luz incandescente (Kréaló, 2013)	25
Ilustración 6. Luz halógena (Arquys, 2012)	26
Ilustración 7. Luz fluorescente (francobordo, s.f.).....	26
Ilustración 8. Luz LED (Almudena, s.f.).....	26
Ilustración 9. Luz Laser (Unisci24, 2014)	27
Ilustración 10. Fibra óptica (CEADE, 2015).....	27
Ilustración 11. Luz frontal (Bcnvision, 2017).....	28
Ilustración 12. Luz lateral (Bcnvision, 2017).....	28
Ilustración 13. Iluminación por campo oscuro (Bcnvision, 2017)	29
Ilustración 14. Iluminación por contraste (Bcnvision, 2017).....	29
Ilustración 15. Iluminación axial difusa (Bcnvision, 2017).....	29
Ilustración 16. Cámara (Definición ABC, s.f.)	30
Ilustración 17. Etapas en la formación de imágenes (Universidad Politécnica de Madrid)	30
Ilustración 18. Lente (FotoNostra, s.f.)	31
Ilustración 19. Sensor CCD (Hisystem, s.f.)	32
Ilustración 20. Sensores CMOS (Terris, 2014).....	32
Ilustración 21. Diagrama básico de las etapas de la visión artificial (Valencia Armijos & Idrovo Urgilés, 2016)	33

Ilustración 22. Logo Python (Rosebrock, 2017)	33
Ilustración 23. Logo OpenCV (Rosebrock, 2017).....	34
Ilustración 24. Categorías de OpenCV (Rodríguez Bazaga, 2015)	34
Ilustración 25. Logo Tensorflow (Inteldig, 2018).....	35
Ilustración 26. DAG (Tensorflow, 2018)	35
Ilustración 27. Representación de rango (Tensorflow, 2018)	36
Ilustración 28. Rango y forma del tensor (Tensorflow, 2018)	36
Ilustración 29. Ejemplo de rango y forma.....	37
Ilustración 30. Histograma de imagen (OpenCV, 2015)	38
Ilustración 31. Imagen con ruido gaussiano (Universidad Politécnica de Madrid, s.f.).....	39
Ilustración 32. Imagen con ruido multiplicativo (Universidad Politécnica de Madrid, s.f.)..	40
Ilustración 33. Imagen con ruido impulso (Universidad Politécnica de Madrid, s.f.).....	40
Ilustración 34. Imagen modelo (MatLab, s.f.)	40
Ilustración 35. Imagen Modelo representada en escala de grises.....	41
Ilustración 36. Imagen modelo / imagen con filtro pasa bajos	41
Ilustración 37. Imagen modelo / imagen con filtro pasa altos	42
Ilustración 38. Imagen modelo/ imagen filtro gausiano	42
Ilustración 39. Imagen modelo con ruido de impulso / imagen modelo con filtro basado en la mediana.....	43
Ilustración 40. Imagen modelo / imagen con filtro bilateral.....	43
Ilustración 41. Bordes (Del Valle Hernández, s.f.).....	44
Ilustración 42. Umbralización (Ipi tel uva, s.f.).....	44
Ilustración 43. Descriptores de frontera por códigos encadenados (Universidad politécnica de Madrid, s.f.).....	45
Ilustración 44. Ejemplo de programar la multiplicación de dos números con Tensorflow .	47
Ilustración 45. Estructura de redes neuronales (Robologs, 2017)	48

Ilustración 46. Redes Convolucionales (GE Intelligent Plataforms).....	49
Ilustración 47. Ejemplo de filtros convolucionales (Yneeds, s.f.)	50
Ilustración 48. Maxpooling vs Averde pooling (Yu, Wang, & Wei, 2014)	50
Ilustración 49. Ejemplo de aumento de datos a partir de una imagen (Raj, s.f.)	51
Ilustración 50. Ejemplo de Dropout (Budhiraja, 2016)	51
Ilustración 51. Pérdida Alta / Pérdida baja (Google, s.f.)	52
Ilustración 52. Pasos muy pequeños / pasos muy grandes (Google, s.f.)	53
Ilustración 53. Orcam MyEye tomado de (Quantumr LV, 2016)	56
Ilustración 54. Eyesynth tomado de (Eyesynth, s.f.).....	56
Ilustración 55. Caja negra	58
Ilustración 56. Diagrama de funciones	59
Ilustración 57. Dimensiones del diseño físico del sistema controlado	64
Ilustración 58. Vista superior del sistema	64
Ilustración 59.vista exterior del sistema controlado	65
Ilustración 60. Vista interna del sistema controlado	65
Ilustración 61. Soporte de cámara para gafas.....	66
Ilustración 62. Vista frontal soporte cámara	66
Ilustración 63. Dimensiones soporte de cámara	66
Ilustración 64. Guante rojo.....	67
Ilustración 65. Guante diseño 2/ Parte palmeada Bolsillo enmarcado.....	67
Ilustración 66. Guante negro de algodón con bordes verdes	68
Ilustración 67. Estructura del código	68
Ilustración 68. Promedio valor de pixel en gris por filas.....	70
Ilustración 69. Promedio del valor de pixel en gris por bloques de filas y columnas.....	71
Ilustración 70. Vecinos del valor central.....	72

Ilustración 71. Patrones local binary paterns.....	72
Ilustración 72. Posición de local binary paterns	72
Ilustración 73. Valor central con nuevo código.....	73
Ilustración 74. Ejemplos de patrones cuando $U=0$	73
Ilustración 75. Ejemplo de patrones cuando $U=2$	74
Ilustración 76. Ejemplo de patrones cuando $U=4$	74
Ilustración 77. Patrones con sus respectivas tipos de características	75
Ilustración 78. Diferentes maneras de separar el conjunto de entrenamiento y el conjunto de validación.....	77
Ilustración 79. Ejemplo validación cruzada	80
Ilustración 80. Foto del sistema controlado	83
Ilustración 81. Imagen color/ imagen gris/ imagen filtrada.....	84
Ilustración 82. Imagen binarizada	84
Ilustración 83. Imagen dilatada / contorno de imagen/ imagen rellena	84
Ilustración 84. Figura rectangular encontrada en imagen binarizada e imagen original ...	85
Ilustración 85. Imagen con perspectiva corregida	85
Ilustración 86. Imagen en gris.....	86
Ilustración 87. Recorrido de las filas en la imagen	86
Ilustración 88. Imagen en gris.....	87
Ilustración 89. Recorrido de promedio por bloques	87
Ilustración 90. Imágenes en cada uno de los canales RGB / histograma de cada una de las imágenes.....	88
Ilustración 91. Imagen HUE / Histograma de imagen En HUE	89
Ilustración 92. Imagen en gris.....	89
Ilustración 93. Histograma en LBP.....	90
Ilustración 94. Imagen color / imagen HSV	94

Ilustración 95. Imagen binarizada	94
Ilustración 96. Recorrido del guante.....	95
Ilustración 97. Imagen a color con los huecos cerrados / imagen a blanco y negro con los huecos cerrados	95
Ilustración 98. Imagen invertida y rellena externamente	96
Ilustración 99. Imagen recortada y combinada con la original.....	96
Ilustración 100: Problemas en el pre procesado de las imágenes del diseño prototipo 1 .	97
Ilustración 101. Imágenes mal segmentadas.....	97
Ilustración 102. Imagen original / imagen posterizada	98
Ilustración 103: imagen posterizada en escala HSV	98
Ilustración 104. Imagen binarizada	99
Ilustración 105. Imagen original recortada con los puntos del contorno interior del guante	99
Ilustración 106. Imagen Pre procesada nuevamente con el código del prototipo 1	99
Ilustración 107: Imágenes de guantes incompletos.....	100
Ilustración 108: Billeto de \$10.000 binarizado.....	100
Ilustración 109. Captura de imagen diseño físico 3.....	101
Ilustración 110. Imagen en gris / imagen binarizada	102
Ilustración 111. Imagen binarizada con rangos del color verde e invertida	102
Ilustración 112. Unión entre las dos imágenes binarizadas.....	103
Ilustración 113. Contorno guante enmarcado en un rectángulo	103
Ilustración 114. Imagen original recortada	103
Ilustración 115. Cambio de tamaño de imagen recortada	104
Ilustración 116. Resumen del diseño de la red neuronal.....	105
Ilustración 117. Disminución rendimiento por número de clases.....	107

LISTA DE ANEXOS

	pág.
Anexo 1. Matriz morfológica	120
Anexo 2. Diseño de procesamiento	123
Anexo 3. La predicción y reconocimiento de imagen	124
Anexo 4. Pre procesamiento experimento 1	125
Anexo 5. Pre procesamiento experimento 2	126
Anexo 6. Pre procesamiento experimento 3	127
Anexo 7. Modelo de deep learnig	128

GLOSARIO

OMS: Organización Mundial de la Salud

INCI: Instituto Nacional para Ciegos

Señalética: consiste en orientar a un individuo dentro de un entorno particular

Braille: un lenguaje para personas discapacitadas visualmente que utiliza una serie de puntos en relieve que se interpretan como letras del alfabeto.

Inclusión Social: proceso que asegura que las personas con discapacidad tengan derecho a la igualdad de oportunidades y a la plena inserción social.

Hogar Taller para Ciegos Ángel de Luz: es una fundación u organización constituidas sin fin de lucro con un carácter de servicio comunitario por la inclusión social de la población con discapacidad visual.

Python: es un lenguaje de programación orientado a objetos, puede distribuirse libremente y Existen versiones para la mayor parte de las plataformas.

Algoritmo: es una secuencia de instrucciones que representan un modelo de solución para determinado tipo de problemas.

Convolución: operación matemática que combina dos señales para producir una tercera señal.

Pata palmeada: patas de animales que tienen los dedos unidos por una membrana por ejemplo, la pata de pato.

Clase: cada objeto que se quiere clasificar o reconocer

Posterizar: Agrupación de los tonos continuos en una imagen.

RESUMEN

Con la entrada en vigencia de la Ley 1618 de 2013 el Estado busca garantizar y asegurar el ejercicio efectivo de los derechos de las personas con discapacidad mediante la adopción de medidas y políticas que cuenten con enfoque de inclusión, lo que permitirá eliminar toda forma de discriminación en el país por cuenta de una condición.

Para el discapacitado visual, esta discriminación se acentúa en las actividades que involucran dinero, debido a que no alcanza una autonomía en la sociedad, por la incertidumbre que hoy en día genera el no contar con una única forma de identificar dos generaciones billetes que circulan en el mercado colombiano, los cuales pierden tanto su marca táctil como el sistema Braille; aunque se creó una plantilla, es un sistema manual basado en el tamaño y funciona únicamente para los billetes de nueva generación. Esto reduce agilidad al discapacitado visual porque incrementan los tiempos en el reconocimiento, arrojando márgenes de error en la identificación del papel moneda, y lo deja en desventaja en las transacciones que realiza.

Para contribuir en este proceso de inclusión se planteó un sistema de reconocimiento de billetes mediante la utilización de las técnicas que ofrece la visión artificial mediante procesamiento de imágenes e inteligencia artificial.

Para la creación del sistema se procesaron las dos generaciones de billetes mediante las siguientes etapas: captura, pre procesamiento, segmentación y reconocimiento involucrando procesos como la remoción de ruido, resaltado de características y optimizaciones.

Se utilizaron algoritmos de inteligencia artificial sobre Python, integrado con librerías como OpenCV, tensorflow y Deep learning.

Como resultado se obtuvieron tres prototipos:

- Cámara endoscópica y caja con ambiente controlado.
- Cámara endoscópica, soporte de cámara para gafas y guante rojo
- Cámara endoscópica, soporte de cámara para gafas y guante negro con borde verde limón brillante.

Finalmente, con los entrenamientos realizados se logró predecir un 96% de exactitud con tiempo de ejecución en tiempo real de 1.5s.

Palabras clave: Reconocimiento, Patrones, Filtro, Ruido, Histograma, Iluminación

ABSTRACT

With the establishment of the Law 1618 of 2013, the State seeks to guarantee and ensure the effective exercise of the rights of people with disabilities through the adoption of measures and policies that have an inclusion focus, which will eliminate all forms of discrimination in the country because of this condition.

For the visually impaired, this discrimination is accentuated in activities that involve money, because these people can not still reach autonomy in society, due to the uncertainty that today generates not having a single way to identify two generations of bills circulating in the Colombian market, which lose both their tactile mark and their Braille system; Although a template was created, it is a manual system based on size and works only for new generation bills. This reduces the agility of the visually impaired because these people need more time to recognize the money, yielding margins of error in the identification of paper money, and leaving them at a disadvantage in the transactions they perform.

To contribute to this inclusion process, a bill recognition system was proposed through the use of techniques offered by artificial vision through image processing and artificial intelligence.

For the creation of the system, the two bill generations were processed through the following stages: capture, pre-processing, segmentation and recognition involving processes such as noise removal, highlighting of features and optimizations.

Artificial intelligence algorithms were used on Python, integrated with libraries such as OpenCV, tensorflow and Deep learning.

As a result, three prototypes were obtained:

- Endoscopic camera and box with a controlled environment.
- Endoscopic camera, camera holder for glasses and red glove
- Endoscopic camera, camera holder for glasses and black glove with bright lemon green border.

Finally, with the training carried out, it was possible to predict a 96% accuracy with real-time execution time of 1.5s.

Keywords: Recognition, Patterns, Filter, Noise, Histogram, Illumination

INTRODUCCIÓN

"Un mundo sin luz sólo se recorre seguro, si en el camino encuentras una mano amiga que te guíe" es el lema del Hogar Taller para Ciegos "Ángel de Luz", fundación escogida para la prueba piloto del presente proyecto que se basa en la visión artificial, campo de la inteligencia artificial que trata de producir el efecto del ojo humano en las máquinas. Éstas pueden percibir y entender una imagen o secuencia de imágenes y actuar según convenga en una determinada situación.

Se puede afirmar que la visión artificial tiene el reto de complementar las insuficiencias visuales, y llegar a ser una herramienta de apoyo para aquellos que no pueden ver.

El proyecto en el primer capítulo se encuentra el planteamiento del problema, objetivo general y marco de referencia:

El planteamiento del problema, expone la situación que los discapacitados enfrentan el día de hoy, donde son varios los desafíos debido a las dos generaciones de billetes que circulan en el mercado colombiano, que resta agilidad al proceso de identificación y dificulta su autonomía en las actividades que involucren dinero, desacelerando el proceso de inclusión social que promueve la ley 1618 de 2013.

El objetivo es desarrollar un sistema de reconocimiento de billetes, para personas con discapacidad visual mediante visión artificial.

El Marco de referencia, se divide en:

- Marco contextual que habla sobre la discapacidad visual en Colombia y la norma 1618 de 2013. Además, todo lo relacionado con billetes de circulación basados en el Banco de la República.
- Marco teórico donde confluyen diferentes temas de autores relacionados con la visión artificial, especificando cada una de las fases de pre procesamiento, segmentación y reconocimiento mediante desarrollo de algoritmos.

En el segundo capítulo se encuentra el método a utilizar, el cual para este proyecto se basó en la metodología de Ulrich, la cual consiste en modelos de fases donde se estructuran el proceso de diseño en etapas y la secuencia a llevar a cabo.

Se espera que el proyecto en una fase posterior pueda alcanzar a desarrollarse de tal manera que permita la identificación de billetes falsos.

1. PRELIMINARES

1.1 PLANTEAMIENTO DEL PROBLEMA

La discapacidad visual en Colombia, representa aproximadamente un 43.5% entre todas las discapacidades existentes en el país; esto equivale al 3% del total de la población (El país, 2013). Esta condición afecta la calidad de vida de una persona, limitando la realización de tareas cotidianas, la interacción con el mundo que lo rodea y con los miembros de su familia; además, demandan mayores recursos en la economía del país, por ser una población, que, en su gran mayoría, requiere apoyo del Estado, por estar al margen de una vida laboral y productiva (OMS, 2014).

En el mundo existen 36 millones de personas que padecen de alto grado de discapacidad visual (OMS, 2017), de las cuales, 1,4 millones pertenecen al territorio colombiano (El país, 2013), siendo las personas de escasos recursos las más afectadas por esta condición. En Colombia, según el periódico El Tiempo, en su artículo “La Falta de Visión no es un Impedimento”, aproximadamente el 80% de los discapacitados visuales sufren pobreza extrema. (El Tiempo, 2015)

La Ley 1618 de 2013, está basada en la inclusión y derechos para las personas con discapacidad, cuyo fin es garantizar la igualdad de oportunidades en todos los ámbitos (Ministerio de Salud y Protección Social, s.f).

En cumplimiento a dicha norma, en materia de inclusión se han implementado las acciones necesarias:

En el transporte masivo

- El metro de Medellín han adecuado las estaciones con ascensores, plataformas mecánicas escaleras eléctricas y/o apoyos isquiáticos (Metro de Medellín, 2018)

El operador y administrador de servicio de transporte público de pasajeros “Solo bus” cuenta con la infraestructura necesaria para permitir el ingreso de usuarios con movilidad reducida. También dispone de sensores sonoros para las personas con dificultades visuales, quienes podrán acceder con perro guía, pues también hay espacio para caninos. Tiene habitáculo para que las personas con discapacidad puedan llevar su silla de ruedas. (Marín Alvarez, 2016)

Estos son avances significativos, pero también son muchos los medios de transporte que carecen de señalética, semáforos sonoros, sistemas de audio y atención (expertos en lenguajes de señas), entre otros (El Tiempo, 2017)

Bibliotecas con formatos alternativos

La Universidad de Antioquia, entre otras instituciones, además de disponer de información académica en formatos alternativos, los usuarios con discapacidad visual cuentan con asesoría personalizada, especializada y permanente para facilitar el logro de sus objetivos académicos. Gracias a esto, 112 personas con discapacidad visual han podido estudiar una carrera profesional completa en la Universidad de Antioquia, graduarse como profesionales y ahora sirven en el medio empresarial público y privado sin que su condición sea inconveniente (Universidad de Antioquia, s.f.)

No obstante, son pocas las bibliotecas que cuentan con formatos alternativos, herramientas pedagógicas y tecnologías de información y comunicación (TIC) para discapacitados.

Reconocimiento de billetes

El Banco de la República ha incorporado señales de sistema braille que permita el reconocimiento de los billetes a los discapacitados visuales. Sin embargo, estos enfrentan varios desafíos debido a las dos generaciones de billetes que circulan en la sociedad colombiana, restando agilidad al proceso de identificación; las siguientes son algunas de las razones:

- De los billetes colombianos en circulación para ser retirados, los elaborados antes del 2010 tienen marca táctil y sistema Braille donde ha perdido su relieve por el desgaste debido al uso frecuente del mismo, dificultando el reconocimiento en forma palpable (véase ilustración 1)



Ilustración 1. Diferencia de braille entre billete de 10.000 nuevo y deteriorado (Banco de la República- Colombia, 2013)

- A partir de 2016, la nueva generación de billetes colombianos en circulación (Banco de la República Colombia, s.f.), también tienen marcas táctiles, que con el tiempo, por el uso tienden a desaparecer y aunque se creó una plantilla medidora de billetes con el sistema braille para identificar su denominación, con base en el tamaño, es una plantilla que:
 - Sólo funciona para los billetes de nueva generación.
 - Es un sistema manual, que reduce la agilidad en cualquier transacción monetaria.
 - No toma en cuenta otras características como colores, textura, contenido, entre otras.

El discapacitado visual debe contar con un ágil sistema de reconocimiento, que unifique la forma de identificación de los billetes en circulación, donde la cantidad de los mismos por contar, la generación y el desgaste del sistema braille, sea irrelevante.

En tiempos modernos, los discapacitados visuales requieren de un sistema que se apoye en la tecnología, permitiendo minimizar los tiempos en las transacciones, reducir los márgenes de error en la identificación del papel moneda y la incertidumbre para los discapacitados visuales; facilitando una mayor autonomía en las actividades que involucren dinero; acercándolos cada vez más a la sociedad colombiana, reduciendo las desigualdades y barreras, dando cumplimiento a lo expresado en la ley 1618 de 2013 (Presidencia de la República, 2013).

Dicho lo anterior, ¿Es posible desarrollar un sistema de reconocimiento billetes, que permita a los discapacitados visuales agilidad en sus transacciones monetarias?

1.2 OBJETIVOS DEL PROYECTO

1.2.1 Objetivo General

Desarrollar un sistema de reconocimiento de billetes para personas con discapacidad visual mediante visión artificial.

1.2.2 Objetivos Específicos

- Definir los billetes de mayor circulación y los patrones característicos de cada uno de ellos.
- Implementar algoritmos de visión artificial para reconocimiento y clasificación de objetos.
- Integrar los algoritmos en un sistema embebido.
- Verificar el funcionamiento del sistema.

1.3 MARCO DE REFERENCIA

1.3.1 Marco Contextual

1.3.1.1 Discapacidad en Colombia

1.3.1.1.1 ¿Qué es discapacidad? *“El resultado de interacciones complejas entre las limitaciones funcionales (físicas, intelectuales o mentales) de la persona y del ambiente social y físico que representan las circunstancias en las que vive esa persona...incluye deficiencias, limitaciones en la actividad y restricciones en la participación, denotando los aspectos negativos de la interacción entre un individuo (con una condición de salud) y la de los factores contextuales individuales (factores ambientales y personales)”* (Ministerio de Salud y Protección Social, s.f)

El National Eye Institute define baja visión como *“una deficiencia visual no corregible por estándar anteojos, lentes de contacto, medicamentos o cirugía que interfiere con la capacidad para realizar actividades cotidianas.”* (News-Medical, 2012)

Según el CDC (Centro para el Control y la Prevención de Enfermedades) y la Organización Mundial de la salud especifican la clasificación visual y deterioro de la siguiente manera (Family connect, s.f)

- La agudeza visual baja visión 20/70 y 20/400
- Ceguera agudeza visual peor 20/400
- Agudeza visual de 20/70 a 20/400

1.3.1.1.2 ¿Cómo es la participación de un discapacitado visual en la sociedad?

El entorno para un discapacitado visual no es acorde a sus necesidades, al no contar con herramientas y señales no visuales suficientes para desenvolverse en el contexto ambiental que le ayuden a suplir aquello con lo que no cuentan (visión) (El Tiempo, 2017) por ejemplo:

- Se carece de dispositivos que informen auditivamente en todos los semáforos el paso de las calles, indicar de manera auditiva los artículos y precios en los supermercados, restaurantes, almacenes de cadena, bibliotecas, entre otros
- Es necesario adecuar la infraestructura de la ciudad, ya que no están señalizados en forma auditiva o sensorial los espacios públicos como parques, centros comerciales, vías, entre otros, con avisos para discapacitados visuales (Observatorio de la Accesibilidad, s.f)
- Se deben adaptar todos medios masivos de transporte, con alertas y avisos de las paradas de buses y estaciones que permitan una ubicación espacial más segura, evitando que se extravíen (Boudeguer Simonetti, Prett Webe, & Squella Fernández, 2010)

- Es fundamental fomentar la literatura física e incentivar el uso de plataformas, aplicaciones, y tecnologías no sólo en bibliotecas y colegios, sino también en las empresas (Ramos Fandiño, s.f).

En el campo laboral, es contada la población discapacitada visual que es productiva, por no contar con los elementos necesarios para su realización personal y profesional; son pocas las empresas que disponen logísticamente de un lugar de trabajo adecuado para ellos. Prefieren contratar a alguien con otro tipo de discapacidad, por ejemplo: con movilidad reducida, sordos, mudos, o con síndrome de down.

Uno de los motivos que impide a los discapacitados visuales laborar es la dificultad para reconocer los billetes, condición necesaria para comprar o vender, ir al banco, tomar un transporte público, entre otros; pues el riesgo de perder dinero es alto, debido a que los billetes generalmente tienen desgastado el sistema braille, situación que deja en desventaja al discapacitado visual en las transacciones que realice.

Finalmente, en el campo emocional y mental, en su gran mayoría son personas aisladas, encerradas en sus casas, con baja autoestima, pues sienten que son una carga, por la dependencia que tienen en su familia y por el costo que requiere su manutención en seguridad social, pensiones, y auxilios cuando se trata de la economía del país. (Ministerio de Educación Nacional, s.f.)

1.3.1.1.3 Ley estatutaria 1618 de 2013

OBJETO. El objeto de la presente ley es garantizar y asegurar el ejercicio efectivo de los derechos de las personas con discapacidad, mediante la adopción de medidas de inclusión, acción afirmativa y de ajustes razonables y eliminando toda forma de discriminación por razón de discapacidad, en concordancia con la Ley 1346 de 2009.

En el Artículo 2, numeral 2: *“Inclusión social: Es un proceso que asegura que todas las personas tengan las mismas oportunidades, y la posibilidad real y efectiva de acceder, participar, relacionarse y disfrutar de un bien, servicio o ambiente, junto con los demás ciudadanos, sin ninguna limitación o restricción por motivo de discapacidad, mediante acciones concretas que ayuden a mejorar la calidad de vida de las personas con discapacidad” (Presidencia de la República, 2013).*

1.3.1.2 Moneda oficial de Colombia

El peso colombiano es la unidad monetaria de curso legal (Gestion Global de Recursos S.L.).

- *Símbolo:* \$
- *ISO:* COP
- *Fecha de creación:* 1810
- *Organismo gestor:* Banco de la República de Colombia

1.3.1.2.1 Papel moneda o billete

“Papel impreso o grabado, generalmente emitido por el banco central de un país, al que se le asigna un valor pecuniario determinado y se emplea como medio legal de pago.” (Oxford, s.f.)

En los últimos tiempos se han diseñado e impreso billetes bajo el sistema braille, para personas que sufren una condición visual permitiendo que reconozcan las diferentes denominaciones, accediendo a ser un usuario del sistema monetario y realizar transacciones que en otro tiempo hubiera sido imposible. Pero cuando el billete está desgastado, deteriorado o defectuoso, se ve afectado dicho reconocimiento y la dependencia e incertidumbre aparece, en el sistema físico implantado en el papel moneda.

1.3.1.2.2 Características de billetes colombianos en circulación

A continuación, se detallan las características de los diferentes billetes colombianos en circulación:

Tabla 1. Billetes colombianos en circulación a partir de 2016 (Banco de la República - Colombia, s.f.)












Características de billetes en circulación a partir de 2016					
Billete (pesos)	Dimensiones	Color Principal	Anverso	Reverso	
\$ 2,000	128 x 66 mm		Azul	Débora Arango	Rio de Caño Cristales
\$ 5,000	133 x 66 mm		Café	José Asunción Silva	Páramos colombianos
\$ 10,000	138 x 66 mm		Rojo	Virginia Gutiérrez	Amazonia colombiana
\$ 20,000	143 x 66 mm		Naranja	Alfonso López Michelsen	Canales la Mojana y Sombrero vueltaio
\$ 50,000	148 x 66 mm		Violeta	Gabriel Gracia Márquez	Ciudad Perdida de la Sierra Nevada
\$ 100,000	153 x 66 mm		Verde claro	Carlos Lleras Restrepo	Valle de Cocora y palma de cera del Quindío

Tabla 2. Billetes colombianos en circulación para ser retirados (Banco de la República - Colombia, s.f.)

Características de billetes en circulación para ser retirados					
---	--	--	--	--	--

Billete (pesos)	Dimensiones	Color Principal		Anverso	Reverso
\$ 1,000	130 x 65 mm		Naranja	Jorge Eliécer Gaitán	Jorge Eliécer Gaitán
\$ 2,000	130 x 65 mm		Beige	Francisco de Paula Santander	Francisco de Paula Santander
\$ 5,000	140 x 70 mm		Verde	José Asunción Silva	José Asunción Silva
\$ 10,000	140 x 70 mm		Café rojizo	Policarpa Salavarrieta	Policarpa Salavarrieta
\$ 20,000	140 x 70 mm		Azul	Julio Garavito Armero	Julio Garavito Armero
\$ 50,000	140 x 70 mm		Púrpura	Jorge Isaacs	Jorge Isaacs

1.3.2 Marco Teórico

1.3.1.1 Sistema de visión humana

La visión humana se define como un sistema integrado entre los ojos, el nervio óptico y el cerebro (Muñoz Manso, 2014). Cuando la luz atraviesa la córnea, la parte externa del ojo, los rayos de luz se desvían, impactando sobre la superficie cristalina, su tarea acaba en la retina, la cual convierte la luz en energía eléctrica que atraviesa el nervio óptico el cual trasmite la señal al cerebro quien se encarga de procesar la información y la manda a través del cuerpo para realizar las diferentes funciones (DocShop, 2014).

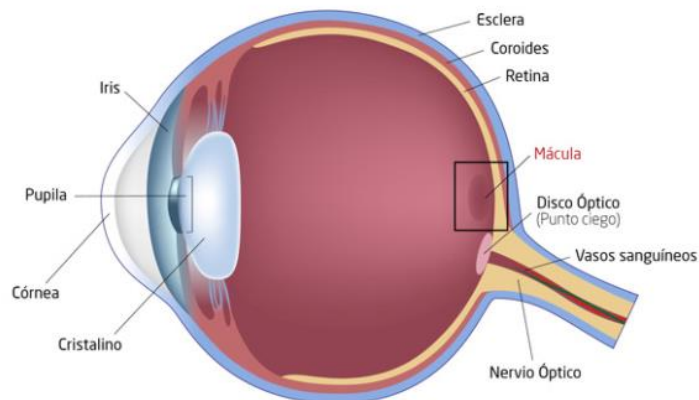


Ilustración 2. Ojo humano (Clínica RAHHAL, s.f.)

Respecto a lo anterior, existen sistemas artificiales que replican dicho sistema integrado; la cámara por ejemplo, permite captar la luz por medio de un lente, el cual pasa por un sensor que traduce la señal lumínica y la convierte en señales eléctricas que recibe el ordenador, y por medio de algoritmos, se puede procesar esas señales. El conjunto integrado del sistema se le conoce como visión artificial.

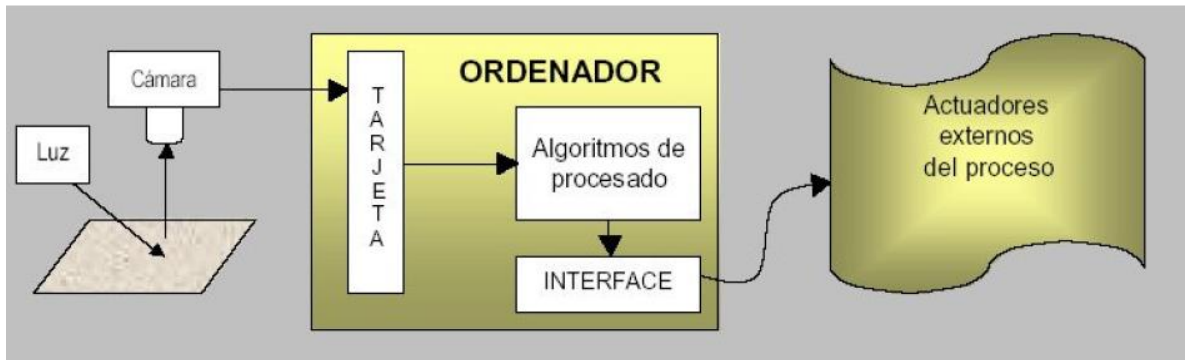


Ilustración 3. Subsistema de las imágenes (Universidad Politécnica de Madrid)

1.3.2.1 Visión artificial

Es conocida como "Visión por Computadora". Conjunto de todas aquellas técnicas y modelos que nos permiten la adquisición, procesamiento, análisis y explicación de cualquier tipo de información espacial del mundo real obtenida a través de imágenes digitales (EcuRed, s.f.).

Vasco Orozco & Tamayo Ocampo la define como un campo de la inteligencia artificial, la cual tiene como finalidad, la extracción de información relevante del entorno exterior a partir de imágenes digitales, mediante la utilización de diferentes técnicas o procesos automáticos que permiten un análisis de esta. (Vasco Orozco & Tamayo Ocampo, 2016).

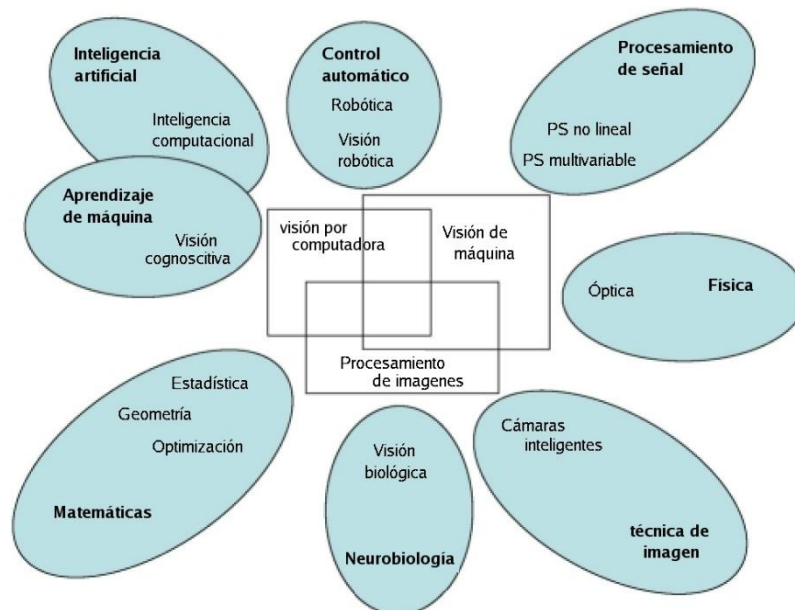


Ilustración 4. Esquema de las relaciones entre visión por computadora y relación (Academic, s.f)

1.3.2.2 Subsistemas de la visión artificial

1.3.2.2.1 Iluminación

Cuando se habla de la visión artificial se hace referencia a la toma de imágenes y a la extracción de características. Una de las variables más importantes a la hora de realizar las capturas es contar con una adecuada iluminación que, creando el ambiente y condiciones necesarias, permita captar una mayor atención visual.

El tipo de iluminación es uno de los factores que más inciden en el sistema, y por tanto necesita de un control exhaustivo, para garantizar la obtención de los resultados deseados a la hora de extraer las características. A continuación, se presentan algunos aspectos de la iluminación.

1.3.2.2.1.1 Tipos de fuentes de luz

El propósito de las fuentes iluminación es adecuar el entorno, para una mejor calidad en la captura de la imagen. Algunos tipos de fuentes de luz son:

- Luz incandescentes: Es la luz tradicional que utilizaban en los hogares y empresas en siglo XX. Se caracteriza por producir iluminación con un tono cálido a bajo costo, pero presenta poca eficiencia y alto consumo de Kw/h, traduce un 15% en luz lumínica y el resto en calor y energía. (Quijada, 2012)

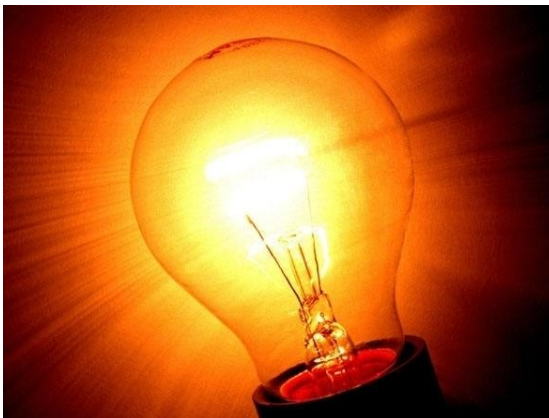


Ilustración 5. Luz incandescente (Kréalo, 2013)

- Luz halógeno: la luz halógena tiene mayor eficiencia, pues es capaz de proporcionar mayor iluminación que el tipo de fuente de luz anterior, puesto que alcanzan potencias mayores. Sin embargo, consumen un alto grado de energía, y presentan riesgo en su manipulación (Quijada, 2012).

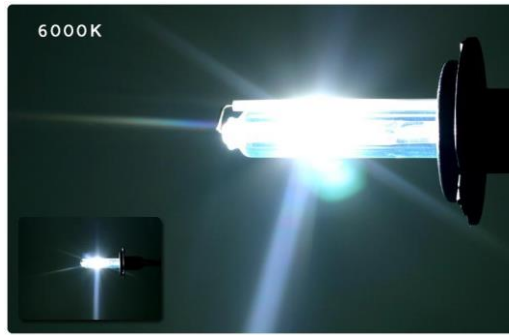


Ilustración 6. Luz halógena (Arquys, 2012)

- Luz fluorescente: las luces fluorescentes tienen extrema eficiencia y contraste, se utilizan en ambientes en donde se necesita luz continua. (Quijada, 2012).



Ilustración 7. Luz fluorescente (francobordo, s.f.)

- Luz LED: las luces LED aportan mucha más energía que cualquiera de las luces, esto lo hacen sin producir calor, lo cual son perfectas para mantener en continuo funcionamiento, ya que consumen menos energía. (Quijada, 2012).



Ilustración 8. Luz LED (Almudena, s.f.)

- Luz láser: La iluminación láser se estructuró para resaltar objetos en tercera dimensión, proyectando una línea de luz en un ángulo conocido, la distorsión de esta permiten una

interpretación en la profundidad del objeto, dependiendo de dónde esté ubicado (Bcnvision, 2017).



Ilustración 9. Luz Laser (Unisci24, 2014)

- Fibra óptica: los haces de fibra óptica proporcionan una iluminación más intensa comparada con los anteriores, conduciendo la luz de una bombilla halógena o de xenón y a través de fibras ópticas, este tipo de luminosidad se debe utilizar en longitudes menores a 5 metros ya que dependiendo de la distancia pierde intensidad de luz. (Infaimon, s.f.)



Ilustración 10. Fibra óptica (CEADE, 2015)

1.3.2.1.2 Técnicas de iluminación

Teniendo presente el tipo de fuente de luz, es necesario un sistema de iluminación, apropiado para los objetos que se quieran reconocer y el entorno dónde debe ir, no todos los sistemas de iluminación están hechos para todas las aplicaciones, es por ello que se

busca encontrar el mejor esquema para que la cámara obtenga mayor información, entre ellos están:

- Luz frontal: La cámara y el haz de luz se posicionan en la misma dirección, directo al objeto, la cámara recibe la luz que refleja el objeto, con el objetivo de eliminar sombras y suavizar texturas. (Bcnvision, 2017)



Ilustración 11. Luz frontal (Bcnvision, 2017)

- Luz lateral: La cámara se posiciona directa al objeto, mientras que el haz de luz se coloca con cierto grado de inclinación dirigido al objeto, esto se realiza para realzar texturas. (Bcnvision, 2017)

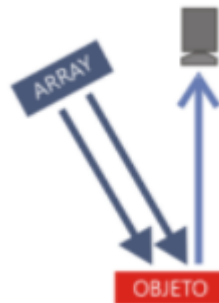


Ilustración 12. Luz lateral (Bcnvision, 2017)

- Darkfield (iluminación por campo oscuro): La cámara se posiciona perpendicular al objeto, mientras el haz de luz se posiciona lateral o alrededor de este, mejorando la visualización de los defectos de las marcas que no se pueden ver a simple vista (Bcnvision, 2017)



Ilustración 13. Iluminación por campo oscuro (Bcnvision, 2017)

- Back light (iluminación por contraste): La cámara se posiciona perpendicular al objeto, mientras el haz de luz debajo de este apunta a la cámara, con el objetivo de detectar imperfecciones, en materiales translucidos (Bcnvision, 2017).



Ilustración 14. Iluminación por contraste (Bcnvision, 2017)

- Iluminación axial difusa: La cámara está dirigida directamente al objeto, mientras que el haz de luz pasa por un espejo que refleja la luz hacia el objeto y el mismo refleja luz hacia la cámara (Bcnvision, 2017).



Ilustración 15. Iluminación axial difusa (Bcnvision, 2017)

Luego de conocer la iluminación adecuada para realizar las características del objeto, como forma, color, textura, entre otras, se procede a capturar la imagen con los usos de cámaras digitales.

1.3.2.2.2 Cámaras digitales

“Una cámara digital es un dispositivo tecnológico capaz de registrar imágenes estáticas o en movimiento de un entorno real y convertirlas a un medio digital”. (Definición ABC, s.f.)



Ilustración 16. Cámara (Definición ABC, s.f.)

1.3.2.2.2.1 Formación de imagen

La captura de una imagen digital se ve reflejada en 4 fases, iniciando con el motivo, el cual es la imagen proyectada del objeto de interés, la cual, por medio de señales lumínicas pasa por un lente generalmente convergente que transmite la luz reflejada por la imagen a un sensor que traduce las señales luminosas recibidas a señales eléctricas que van directo al computador como se puede apreciar en la ilustración 17 (Universidad Politécnica de Madrid).

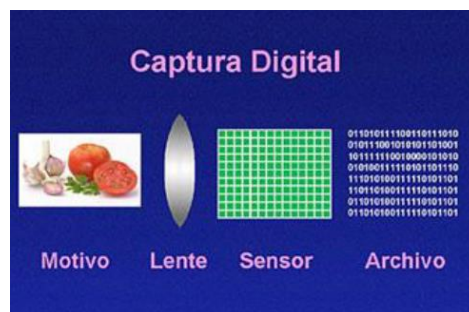


Ilustración 17. Etapas en la formación de imágenes (Universidad Politécnica de Madrid)

1.3.2.2.2.2 Lente

El lente de la cámara permite el paso de los rayos de luz y el enfoque para formar una imagen, esto lo hace por medio de formas cóncavas del lente que permiten que los rayos converjan hasta el punto focal, formando una distancia entre el centro óptico del lente y el punto focal, que en este caso sería el sensor, llamada la distancia focal (FotoNostra, s.f.), esto se puede apreciar en la siguiente figura.

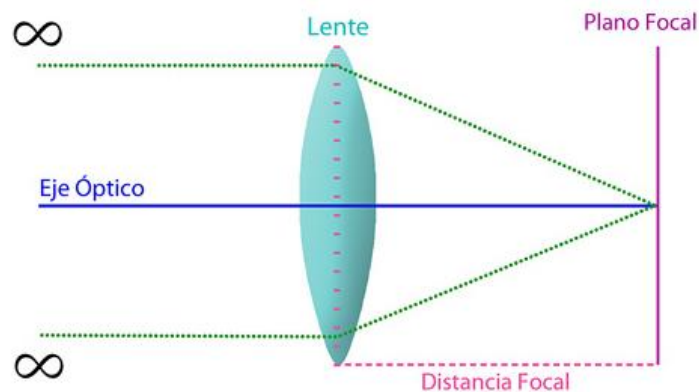


Ilustración 18. Lente (FotoNostra, s.f.)

Dos aspectos importantes a considerar de un lente es su distancia focal, ya que determina la ampliación de la imagen (que tanto de la escena se vea) y el diafragma, la cual indica la apertura del lente la cual permite el paso de cierta cantidad de luz. En algunas cámaras estas dos características suelen ser variables (García, 2018).

1.3.2.2.3 Sensores

Un sensor es un dispositivo que está capacitado para detectar acciones o estímulos externos y responder en consecuencia. Estos aparatos pueden transformar las magnitudes físicas o químicas en magnitudes eléctricas, existe una gran variedad de sensores para un sinfín de aplicaciones, por ejemplo sensores capacitivos, sensores de posición, de presión y fuerza, sensores láser, y sensores ópticos, entre otros (Mecalux, s.f.).

Los sensores de la cámara funcionan como transductores de una señal de luz a una señal eléctrica que va directo a la tarjeta, la cual se encarga de convertir esas señales eléctricas en señales binarias que pueda leer el computador, los más comunes son:

- CCD: es uno de los sensores más utilizados en el mercado, ya que presenta muy buena calidad de imagen y muy buena sensibilidad a la luz, este lo hace por medio de celdas solares estructuradas en forma de matriz (filas y columnas) que convierten la intensidad lumínica que incide sobre cada una de ellas, a una señal análoga formando una porción de la imagen (pixel), luego estas señales analógicas se traducen en una señal binaria, por medio de un convertidor ADC acoplado, que posteriormente forma el arreglo de la imagen. Estos sensores se caracterizan por su buena resolución y la disminución ruido en la imagen, sin embargo consumen mayor energía (Palazzesi, 2006).

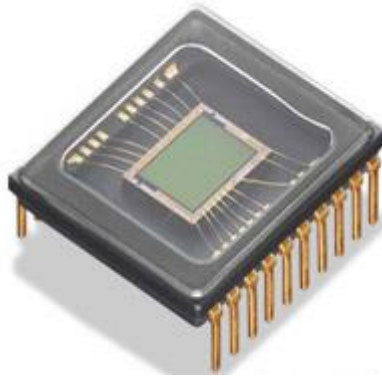


Ilustración 19. Sensor CCD (Hisystem, s.f.)

- CMOS: es un sensor de imagen que posee una calidad un poco más reducida que el anterior, ya que las celdas son independientes la una de la otra y la digitalización se realiza pixel a pixel dentro del sensor, sin tener que acoplar otro circuito, y la salida de este es binaria, lo que reduce el tiempo de procesamiento de la información, para convertir la imagen en una matriz binaria. cada celda está conformada por transistores los cuales amplifican la señal de luz y procesa la información, todo dentro del mismo sensor. Tiene mayor velocidad de captura de imagen, bajo costo y un consumo de energía menor; sin embargo, no reduce tanto el ruido y su sensibilidad es muy baja. (Palazzesi, 2006).

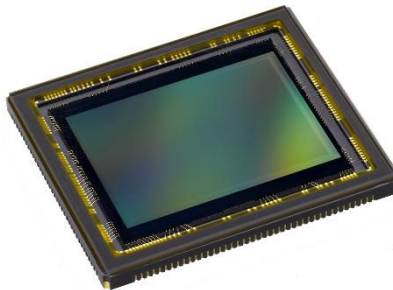


Ilustración 20. Sensores CMOS (Terris, 2014)

1.3.2.2.3 Ordenador y procesador de imágenes

Continuando con el esquema de la imagen 3, después de obtener la representación bidimensional de la escena (la captura de la imagen), se procede a realizar un procesamiento como se puede apreciar en la siguiente figura:

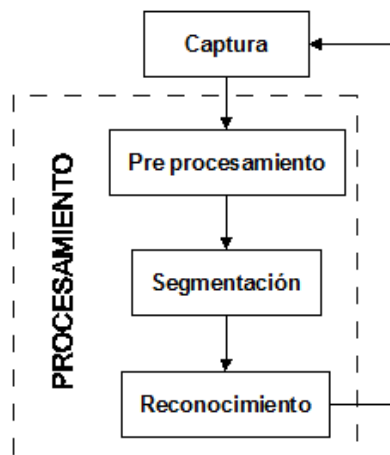


Ilustración 21. Diagrama básico de las etapas de la visión artificial (Valencia Armijos & Idrovo Urgilés, 2016)

En donde involucra:

- *Pre procesamiento*: permite mejorar o realzar las propiedades para optimizar las etapas siguientes, aquí es donde se aumenta el contraste, suaviza o elimina el ruido y detecta bordes, mediante el uso de diferentes técnicas informáticas. (Universidad politécnica de Madrid, s.f.)
- *Segmentación*: la segmentación permite obtener las regiones de interés a analizar y sus características. (Universidad Politecnica de Madrid)
- *Reconocimiento*: la finalidad del reconocimiento es distinguir objetos segmentados, gracias al análisis de características que previamente se han establecido, para diferenciarlos (Universidad politécnica de Madrid, s.f.).

1.3.2.2.3.1 Software

El software que reúne mayores oportunidades de realizar diferentes tipos de aplicaciones es Python, el cual es un software de código libre diseñado para multiplataforma que permite el uso de diferentes librerías para la realización de diversas aplicaciones.



Ilustración 22. Logo Python (Rosebrock, 2017)

En cuanto a los recursos para el procesamiento de imágenes, se han creado librerías como OpenCV y Tensorflow, las cuales son consideradas como las mejores y más completas

para su especialidad. Actualmente empresas reconocidas como Google o Microsoft, utilizan estas herramientas para sus aplicaciones en visión artificial (OpenCV, s.f.)

1.3.2.2.3.1.1 Open CV



Ilustración 23. Logo OpenCV (Rosebrock, 2017)

Es una librería de visión por computadora y aprendizaje automático, que fue diseñado principalmente para optimizar algoritmos (el uso de recursos computacionales) de visión artificial y acelerar el uso de la percepción de la máquina, Con OpenCV se han realizado diferentes aplicaciones en diferentes áreas, como en seguridad, medicina, sector industrial, entre otros. (OpenCV, s.f.)

OpenCV es una librería de código libre, escrito en C++, y cuenta con interfaces para C, Python, Java, Matlab; Adicional a esto funciona en múltiples sistemas operativos como Windows, Android, Linux, Mac Os. (OpenCV, s.f.)

Además permite visualizar datos y extraer información de la imagen y/o videos, contando con diferentes categorías y formas de realizar esto como, realizar filtros, segmentar objetos, realizar algoritmos de inteligencia artificial, detectar objetos, encontrar características, entre otros; esto se puede ver en la siguiente figura (Rodríguez Bazaga, 2015)



Ilustración 24. Categorías de OpenCV (Rodríguez Bazaga, 2015)

1.3.2.2.3.1.2 Tensorflow



Ilustración 25. Logo Tensorflow (Inteldig, 2018)

Es una librería de código abierto diseñada para realizar cálculos numéricos de manera más eficiente y de alto rendimiento, mediante gráficos de flujos de datos, el cual se les conoce como Gráficos Acíclicos Dirigidos (DAG), este gráfico se construye a partir de nodos y bordes (Google Cloud, s.f.).

- **Los nodos** del gráfico representan las operaciones matemáticas (sumas, restas, multiplicación, entre otras) y funciones como softmax, sigmoid, entre otras. (Google Cloud, s.f.).
- **Los bordes** del gráfico representan las conexiones entre los nodos, los cuales son las entradas y salidas de las operaciones matemáticas (nodos). (Google Cloud, s.f.).

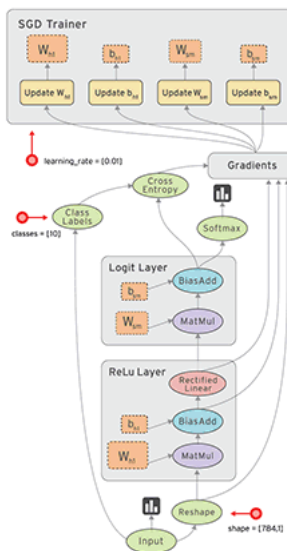


Ilustración 26. DAG (Tensorflow, 2018)

Además, es una librería muy utilizada para este tipo de desarrollos (visión por computadora), y fue diseñada por Google como una librería de código abierto, presenta muy buenos recursos a la hora de manejarla, permitiendo no sólo realizar aplicaciones con imagen, sino con cualquier tipo de datos como voz, texto, entre otros, mediante datos llamados tensores. (Google Cloud, s.f.)

Los tensores representan los datos en forma de matrices multidimensionales que fluyen a través de la DAG, estos se componen de un rango y una forma.

- **Rango:** número de dimensiones del tensor (Tensorflow, 2018)

Rank	Math entity
0	Scalar (magnitude only)
1	Vector (magnitude and direction)
2	Matrix (table of numbers)
3	3-Tensor (cube of numbers)
n	n-Tensor (you get the idea)

Ilustración 27. Representación de rango (Tensorflow, 2018)

- **Forma:** el número de elementos que hay en cada dimensión (Tensorflow, 2018)

Rank	Shape	Dimension number	Example
0	[]	0-D	A 0-D tensor. A scalar.
1	[D0]	1-D	A 1-D tensor with shape [5].
2	[D0, D1]	2-D	A 2-D tensor with shape [3, 4].
3	[D0, D1, D2]	3-D	A 3-D tensor with shape [1, 4, 3].
n	[D0, D1, ... Dn-1]	n-D	A tensor with shape [D0, D1, ... Dn-1].

Ilustración 28. Rango y forma del tensor (Tensorflow, 2018)

Para aplicaciones con imágenes se utilizan rangos a partir de dos dimensiones, en donde se tienen como valores iniciales el tamaño de la imagen, a medida que aumenta los valores, como el canal o el número de lotes, el tensor aumenta su dimensión, como se muestra en la siguiente figura:



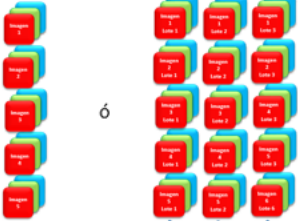
Rango	Tipo de imagen	Matriz	Forma	Representación en imagen
2	Imagen en escala de grises	[Ancho, Alto]	[224, 224]	
3	Imagen a Color	[Ancho, Alto, Canal]	[224, 224, 3]	
4	Lote de imágenes para uno o varios objetos	[Cantidad imágenes, Ancho, Alto, canal]	[5, 224, 224, 3] o [15, 224, 224, 3]	

Ilustración 29. Ejemplo de rango y forma

Tensorflow es utilizado mayormente para el marco del aprendizaje profundo (Deep learning), el cual permite extraer características mediante procesos jerárquicos, en donde las capas iniciales de la jerarquía aprenden características simples, que luego envía la información a las siguientes capas, que realizan combinaciones para ir formando partes del objeto cada vez más complejas, finalizando con la jerarquía superior, en donde aprende a clasificar (García Moreno, 2016)

Si bien la librería anterior puede realizar acciones similares, con tensorflow se puede obtener mejores resultados, tanto en el rendimiento como en la precisión a la hora de realizar los cálculos de entrenamiento, sin necesidad de tener que realizar un pre procesamiento extenso, puesto que permite realizar técnicas mejores como lo es aprendizaje profundo (García Moreno, s.f.).

1.3.2.2.4 Imagen digital

Entiéndase como imagen digital:

“la representación bidimensional de una imagen empleando bits, unidad mínima de información compuesta por dígitos binarios (1 y 0), que se emplea a instancias de la informática y cualquier dispositivo de tipo digital” (Definición ABC, s.f.).

Algunos elementos que se deben tener en cuenta a la hora de trabajar con imágenes digitales son. (Gomez, 2015):

- Pixel: Es el mínimo valor que compone una imagen.
- Número de canales: es la representación visual que interpretan las computadoras, Por ejemplo, si la imagen tiene un solo canal significa que está en escala de grises; si la imagen cuenta con más canales se puede apreciar la imagen a color.
- Espacio de color: la gama de colores que representa el computador (Musso, 2018), el espacio de color que captura las cámaras es RGB (Red, Green, Blue), a partir de este espacio de color se realizan modificaciones a otros espacios como HSV, HLS, CMYK. Cada uno de ellos contienen valores representados por matrices de imágenes conocidas como canales.
- Resolución: nivel de detalle con el que puede observarse la imagen
- Tamaño: la cantidad de píxeles que existen en una imagen (fila X columna)

1.3.2.2.5 Pre procesamiento

Esta etapa de la visión artificial suele ser una de las más importantes, en donde se implementan algoritmos como:

1.3.2.2.5.1 Histograma

El histograma de una imagen en grises permite contabilizar el número píxeles de cada nivel, representándose en un eje de coordenadas (Universidad Politécnica de Madrid, s.f.) como se puede ver en la siguiente figura.

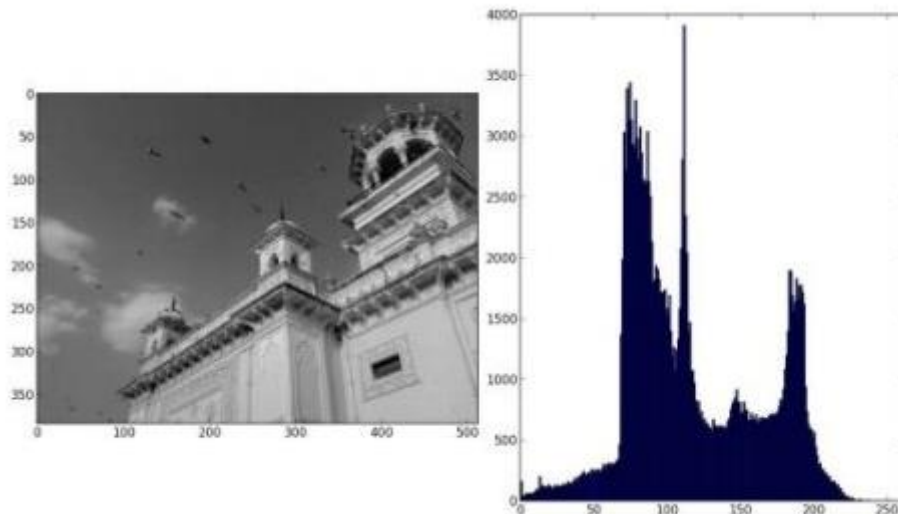


Ilustración 30. Histograma de imagen (OpenCV, 2015)

En donde se puede analizar, el contraste y la saturación

- El contraste: se encarga de medir la dispersión de los niveles de grises en la imagen, entre menor contraste haya, menor variabilidad de la intensidad (niveles de grises) hay en la imagen (Universidad Politécnica de Madrid, s.f.).
- Saturación: cuando se habla de saturación, se hace referencia a que el histograma de la imagen presenta valores muy altos en sus extremos (Universidad Politécnica de Madrid, s.f.).

1.3.2.2.5.2 Ruido

La imagen en la mayoría de las ocasiones presenta ruido que degrada la imagen, esto es lo primero que debe si no bien eliminar, entonces reducir al máximo, pero antes, para poder reducir el ruido, se debe conocer que ruido es el que se presenta en la imagen, como:

1.3.2.2.5.2.1 Tipos de ruido

- Ruido Aditivo o gaussiano: Es un ruido que se presenta a la hora de capturar la imagen, debido al proceso de convertir los fotones en carga eléctrica, este ruido se puede observar un ruido blanco. (Universidad Politécnica de Madrid, s.f.)



Ilustración 31. Imagen con ruido gaussiano (Universidad Politécnica de Madrid, s.f.)

- Ruido Multiplicativo: Es un ruido que se genera por falta de iluminación uniforme, su nombre es debido a la forma en la que se ve, es como se hubieran dos imágenes multiplicadas, una con una intensidad de luz sobre el objeto y la otra con una textura de los objetos de la escena. (Universidad Politécnica de Madrid, s.f.)

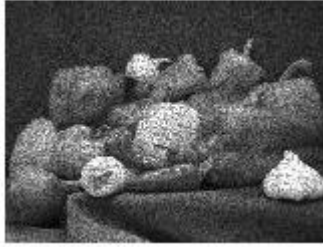


Ilustración 32. Imagen con ruido multiplicativo (Universidad Politécnica de Madrid, s.f.)

- Ruido Impulso: Es un ruido que se genera debido a la saturación de la carga que recibe un pixel, como una perturbación ya sea por exceso o por defecto, este ruido se conoce mejor como, ruido de “sal y pimienta” (Universidad Politécnica de Madrid, s.f.).

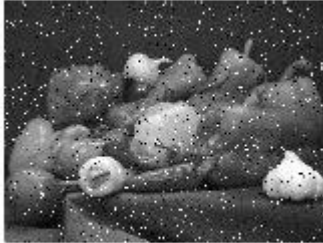


Ilustración 33. Imagen con ruido impulso (Universidad Politécnica de Madrid, s.f.)

1.3.2.5.2.2 Técnicas de suavizado

A partir de las imágenes con diferentes ruidos extraídas (Universidad politécnica de Madrid, s.f.) la autora procedió a realizar técnicas para mostrar los diferentes procesos de reducción de ruido.



Ilustración 34. Imagen modelo (MatLab, s.f.)



Ilustración 35. Imagen Modelo representada en escala de grises

A partir de la imagen a escala de grises se procede a realizar los diferentes filtros

- Filtro pasa baja: se utiliza para reducir el ruido de frecuencias altas, sin embargo, tiende a desvanecer los bordes y los detalles finos en la imagen (Universidad politécnica de Madrid, s.f.)

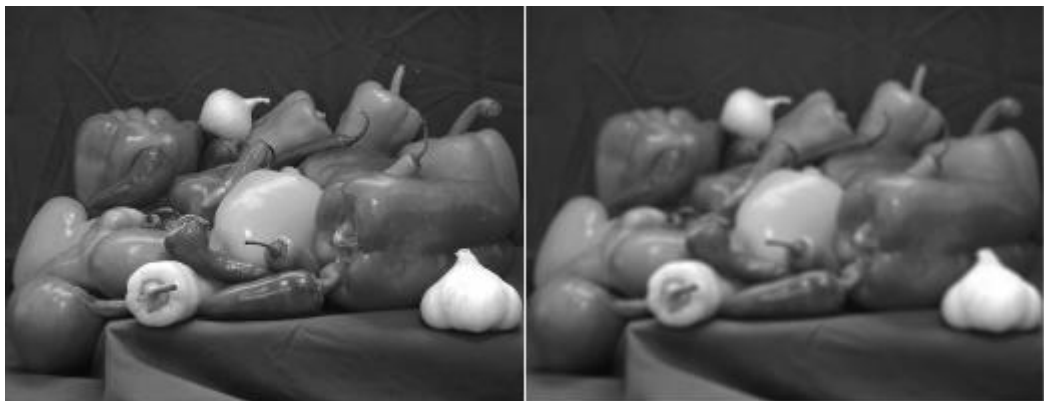


Ilustración 36. Imagen modelo / imagen con filtro pasa bajos

- Filtro pasa alta: hacen el efecto contrario, ya que suelen reducir el ruido de frecuencias bajas, se utilizan para recuperar detalles perdidos, o detalles finos.



Ilustración 37. Imagen modelo / imagen con filtro pasa altos

- Filtro gaussiano: utilizan máscaras de convolución de filtros pasa bajos cuyos anchos de banda depende de sus varianzas, estas máscaras recorren toda la imagen reduciendo ruidos gaussianos. (Universidad politécnica de Madrid, s.f.)

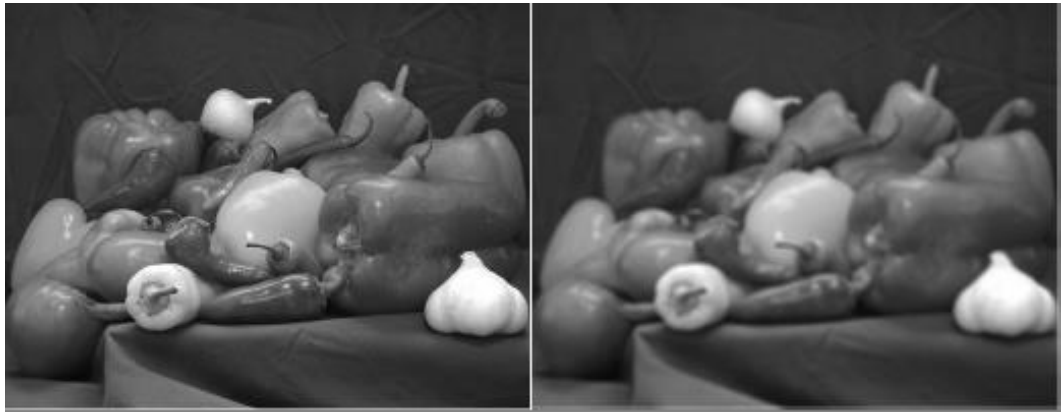


Ilustración 38. Imagen modelo/ imagen filtro gausiano

- Filtro basado en la mediana: este tipo de filtro, elige un valor de mediana de los pixeles más cercanos, y los reemplaza para obtener un mejor suavizado y, este filtro se caracteriza por reducir hasta eliminar el ruido impulsional o de sal y pimienta. (Universidad politécnica de Madrid, s.f.)



Ilustración 39. Imagen modelo con ruido de impulso / imagen modelo con filtro basado en la mediana

- Filtro bilateral: Es un tipo de filtro que conserva los bordes, eligiendo un valor en base a una media ponderada de los píxeles vecinos, el filtro actúa eliminando solo pequeñas diferencias de los píxeles en la imagen.

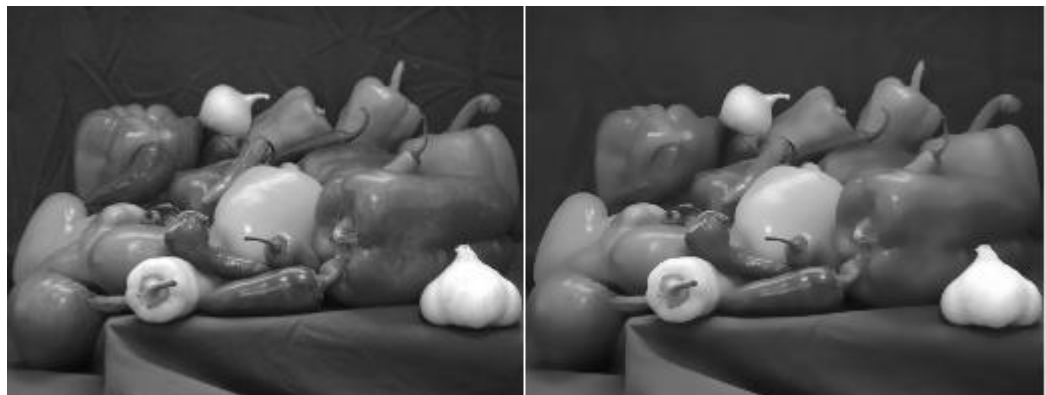


Ilustración 40. Imagen modelo / imagen con filtro bilateral

1.3.2.2.6 Segmentación

Luego de realizar el pre procesamiento, se halla la región de interés contenida en la imagen para su posterior análisis y extracción de características. Existen diferentes técnicas de segmentación:

- Técnicas basadas en bordes: Esta técnica pretende encontrar formas geométricas como líneas rectas y/o curvas a partir de imágenes binarizadas, al tener contornos o bordes representados en una imagen que dependen de una dirección y un módulo de gradiente similar entre píxeles (Universidad politécnica de Madrid, s.f.)



Ilustración 41. Bordes (Del Valle Hernández, s.f.)

- Umbralización: Esta técnica se utiliza cuando el fondo de la imagen se diferencia del objeto a analizar; a partir de una imagen en grises el valor de umbral permite binarizar convirtiendo los valores de los píxeles en blanco, y el resto en negro. El umbral se elige por medio de los valores del histograma en grises, analizando los picos máximos y los valles que tienen, con el fin de obtener una mejor binarización, (Universidad Politécnica de Madrid, s.f.)



Ilustración 42. Umbralización (Ipi tel uva, s.f.)

- Segmentación orientada por regiones: Esta técnica se divide en:
 - La primera técnica es por crecimiento de regiones, que consiste en aumentar el tamaño de cada píxel que corresponda a la propiedad de la conectividad de los píxeles, analizando sus vecinos, según la regla de similitud de la región,
 - La segunda, es la técnica de las pirámides y árboles cuaternos, en donde la imagen se divide en cuatro rectángulos iguales y cada región se vuelve a dividir en otras cuatro regiones, y así sucesivamente hasta alcanzar a nivel de píxel, promediando los píxeles en reducción de la resolución espacial de un nivel a otro.
 - Otra técnica a analizar es la de división y fusión de regiones, en donde se descompone la imagen en regiones no uniformes y si la región suele tener

relación con el contraste o la varianza de la misma se unifica con las otras, formando técnicas de crecimiento de regiones, de lo contrario se dividirá la región (Universidad politécnica de Madrid, s.f.)

1.3.2.2.7 Interpretación de la imagen

Luego de realizar técnicas de segmentación y delimitar los objetos de interés del fondo, se procede a extraer características relevantes según lo especifica la universidad politécnica de Madrid:

- Medidas geométricas: son medidas que están relacionadas con la distancia euclidiana entre pixeles como:
 - El Área: la cantidad de pixeles que hay dentro de un objeto
 - El perímetro: la longitud del borde o contorno del objeto
 - El centro de gravedad o centroide: la posición puntual en la imagen
- Descriptores topológicos: las propiedades topológicas describen el objeto de forma invariante a rotaciones, traslaciones o cambios de escala como:
 - Relación aspecto: la relación que hay entre la altura y el ancho del objeto
 - Cantidad agujeros: la cantidad de agujeros internos al objeto
 - Número de Euler: la diferencia entre componentes conectados de la región menos el número de agujeros de ésta.
- Textura: se define como la composición superficial y visual invariante de un objeto, ya sea por su rugosidad, granularidad, regularidad, entre otras.
- Descriptores de frontera por códigos encadenados: A partir del contorno del objeto, se determina la conectividad, partiendo de un origen determinado y continuando con un sentido horario o anti horario, que codifica la frontera del objeto.

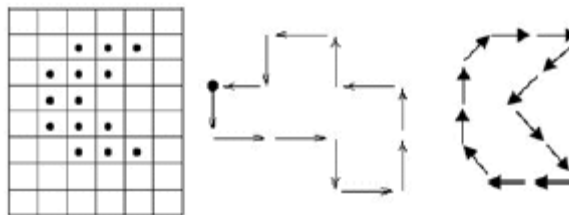


Ilustración 43. Descriptores de frontera por códigos encadenados (Universidad politécnica de Madrid, s.f.)

- Descriptores geométricos: se suele utilizar cuando existen regiones cerradas, en donde utiliza un punto interior del objeto (usualmente el centroide), y codifica la distancia entre el punto interior y la frontera o borde, para luego normalizarla y elegir una distancia máxima entre 0 y 1.

1.3.2.2.8 Reconocimiento de imágenes

El aprendizaje automático (machine Learning), es una disciplina científica del ámbito de la inteligencia artificial que ha tomado auge en últimos tiempos especialmente para reconocimiento de imágenes, permitiendo el uso de los métodos que desarrolla algoritmos que permiten que las máquinas puedan aprender por sí mismas. (Imnovation, s.f.)

Existen diferentes métodos para el aprendizaje automático: supervisados y no supervisados. Los sistemas supervisados consisten en emplear un conjunto de datos conocidos (conjunto de datos de entrenamiento), esto se compone de datos de entrada (Características) y valores de respuesta (etiquetas previamente asignadas a cada objeto); a partir de esto, el algoritmo de aprendizaje supervisado busca crear un modelo que pueda realizar predicciones acerca de los valores de respuesta para un nuevo conjunto de datos. (MathWorks, s.f.)

Para estos métodos existen diferentes algoritmos como: máquina de vectores de soporte (SVM), redes neuronales, clasificador de Naive Bayes, árboles de decisión, entre otros (MathWorks, s.f.).

A diferencia del aprendizaje supervisado, Al aprendizaje no supervisado, sólo se le proporcionan los datos de entrada (características), sin asignarles ningún valor de respuesta (etiqueta), la función del algoritmo es relacionar y crear grupos de datos dependiendo de la similitud de estos; entre más datos se les proporcione, mejor será el rendimiento del algoritmo y el agrupamiento de las características. (Zambrano)

Para estos métodos existe algoritmos como: clustering, análisis de componentes principales (PCA), análisis de componentes independientes (ICA), entre otros. (Raona, 2017)

Para el reconocimiento de imágenes existen varias librerías como:

Scikit-learn

Scikit-learn es una de las librerías más utilizadas en el campo de la inteligencia artificial, herramienta simple y eficiente para la minería y/o análisis de datos, permitiendo así diseñar modelos y entrenarlos de manera muy simple (scikit-learn, s.f.).

Inicialmente se cargan los datos que se quieren reconocer, luego se elige el tipo de modelo que se quiere utilizar, y finalmente se configuran los hiperparámetros con el fin de mejorar el modelo.

Existe una gran cantidad de modelos, dependiendo de la aplicación que se quiera ya sea para clasificar, para aproximar una curva, entre otras, se elije el modelo adecuado.

- **Para clasificar:** Perceptron multicapa (MLP), Support vector machine (SVM), Naive Bayes, regresión logística, entre otros (scikit-learn, s.f.)
- **Para regresión:** Regresión lineal, Support vector Regression (SVR), Regresión Isotónica, entre otros (scikit-learn, s.f.)

Tensorflow

Tensorflow utiliza una arquitectura modular que se conoce como evaluación perezosa, ya que permite una flexibilidad y optimización en la medida que se ejecuta el gráfico, además de lo anterior Tensorflow pretende optimizar los procesos pudiéndose implementar un código en varios computadores a la vez, es así pudiendo distribuir las tareas y agilizar los procesos, utilizando múltiples GPU (Google Cloud, s.f.).

La creación y ejecución de un programa en Tensorflow, es muy diferente en la forma de programarse, se realiza de la siguiente manera:

1. Definir tipo de variable
2. Crear gráfico
3. Ejecutarlo en una sesión
4. Evaluar el tensor en la sesión

Un ejemplo simple de cómo se programa la multiplicación de 2 números, se puede ver a continuación.

```
8 import tensorflow as tf
9
10 #Construir el Gráfico
11 a = tf.constant(5.0) #Definir tipo de variable
12 b = tf.constant(6.0) #Definir tipo de variable
13 c = a * b
14
15 #Ejecutar el Gráfico en una sesión
16 sess = tf.Session()
17
18 #Evaluar el tensor en la sesión
19 print (sess.run(c))
```

Ilustración 44. Ejemplo de programar la multiplicación de dos números con Tensorflow

Keras

Keras es una API de redes neuronales de alto nivel capaz de ejecutarse sobre Tensorflow, CNTK o Theano (Keras, s.f.), hecha con el objetivo de realizar experimentaciones rápidas. Keras se caracteriza por su:

- Facilidad de uso: fácil de comprender y más intuitivo que código escrito directamente en Tensorflow
- Modularidad: forma modelos mediante secuencias de módulos independientes que pueden ser configurados.
- Facilidad de extensibilidad: módulos adicionales son fáciles de agregar
- Trabajar con Python: es un código fácil de depurar y permite extensibilidad.

Ahora para diseñar y crear un modelo de redes neuronales en keras es más sencillo porque únicamente:

1. Se define el modelo
2. Se compila el modelo
3. Se entrena
4. Se predice

- **Diseño de Modelos**

Todos los modelos diseñados en Keras deben seguir una secuencia en donde le indica la capa de entrada, las capas ocultas que le siguen y finalmente la capa de salida, como se puede ver en la siguiente figura.

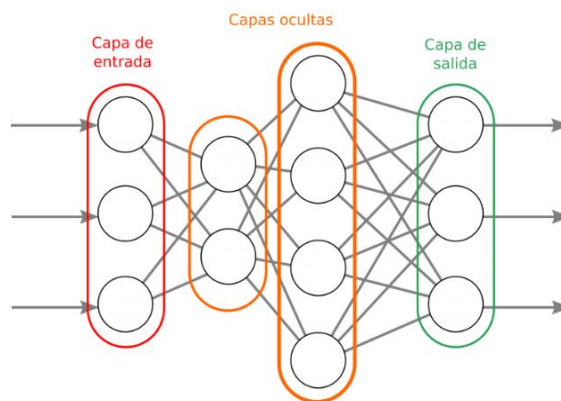


Ilustración 45. Estructura de redes neuronales (Robologs, 2017)

Uno de los mejores modelos para trabajar con imágenes se llaman modelo de redes convolucionales (CNN), que permiten extraer información de la imágenes sin tener que realizar pre procesamiento previo, ya que realizando filtros convolucionales extrae características como texturas, formas, colores, entre otras, esto lo realiza mediante valores de píxeles de la imagen original y aprendiendo cómo extraer estas características para así realizar una correcta clasificación (Google , s.f.).

Las CNN, son modelos de aprendizaje profundo que funcionan con niveles jerárquicos como se mencionó anteriormente, en donde las capas iniciales de la jerarquía aprenden características simples, que luego envían la información a las siguientes capas, que realizan combinaciones para ir formando partes del objeto cada vez más complejas, finalizando con la jerarquía superior, en donde aprende a clasificar (Google , s.f.).

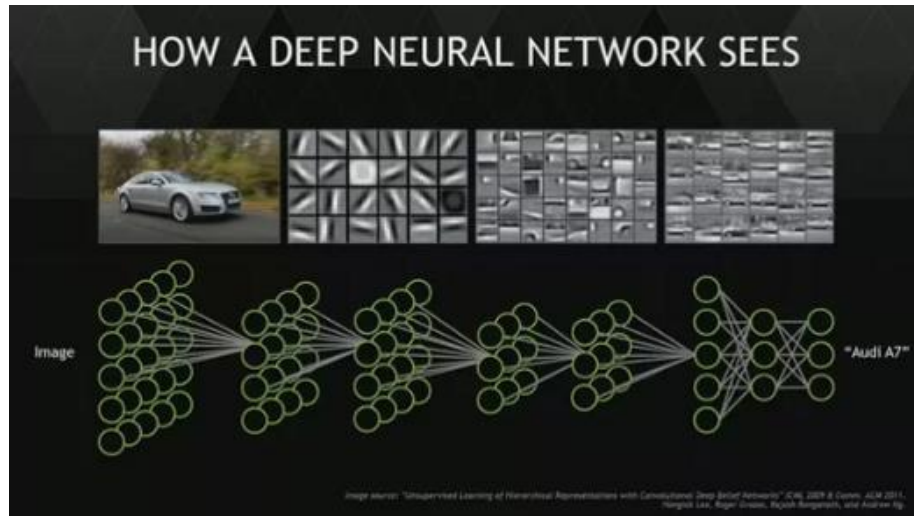


Ilustración 46. Redes Convolucionales (GE Intelligent Plataforms)

Se puede apreciar que las capas iniciales extraen las características de bordes y polígonos de la imagen, las capas siguientes realizan combinaciones de las características anteriores para formar figuras con mayor complejidad y que también se combinan para obtener la imagen final aprendiendo a identificar la.

Algunos términos que se debe tener en cuenta a la hora de realizar algoritmos de inteligencia artificial en keras son:

- **Entradas:** Los modelos CNN reciben como entradas tensores de rango 4, como se vio anteriormente, se componen de la cantidad de imágenes, longitud, ancho y canales y deben entrar a la red con el mismo tamaño (Google , s.f.).
- **Filtros Convolucionales (Conv 2D):** Los modelos realizan convoluciones aplicando filtros que permiten extraer características como textura, bordes, formas, entre otros, que dan como salida un mapa de características, entre mayor sea el número de filtros (profundidad de la salida del mapa), mayor es el tiempo de entrenamiento y por ende mayor el consumo de la máquina. Se recomienda tener un mínimo de filtros necesarios que realice una clasificación precisa (Google , s.f.)

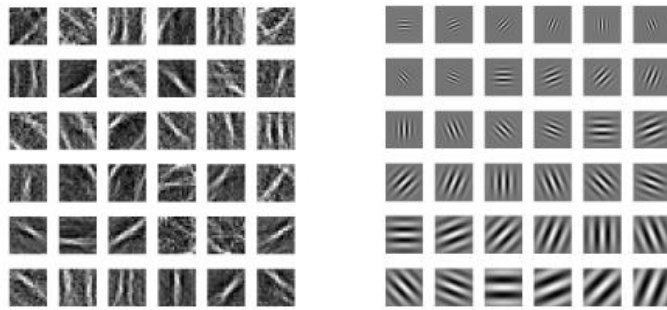


Ilustración 47. Ejemplo de filtros convolucionales (Yneeds, s.f.)

- **Función de activación (Activation):** La función de activación, debe introducir la no linealidad al sistema, aplicando transformaciones no lineales a las características convolucionadas, la más utilizada es la unidad lineal rectificadora (ReLU), la cual devuelve el valor de la característica siempre y cuando los valores sean positivos (Google, s.f.)
- **Normalización por lotes (Batch normalization):** Después de la función de activación, es importante normalizar los datos, además de las entradas también es necesario normalizar las salidas de las capas ocultas, permitiendo acelerar el entrenamiento, ya que los datos estarían en un rango entre 0 y 1. Este paso suele ser opcional (Doukkali, 2017).
- **Agrupación Máxima (Maxpooling):** Después de realizar los pasos anteriores, lo siguiente es el agrupamiento, con el objetivo de reducir el número de dimensiones del mapa de características conservando la información de características más importante, esto lo realiza con un kernel que recorre la imagen, encontrando el valor máximo dentro del cada deslizamiento por la imagen, enviándose ese valor máximo a un nuevo mapa de características, esto suele reducir el mapa de características a la mitad (Google Cloud, s.f.)

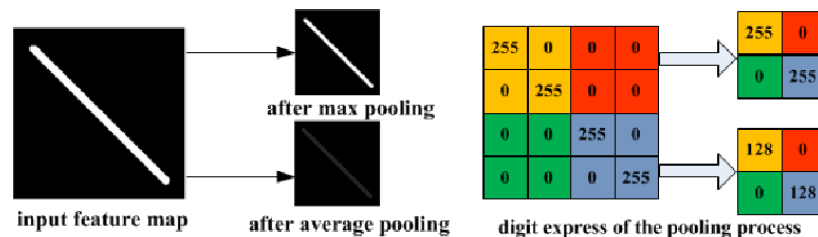


Ilustración 48. Maxpooling vs Average pooling (Yu, Wang, & Wei, 2014)

- Para evitar o prevenir el sobreajuste y que la red sólo aprenda datos de entrenamiento existen varias técnicas:
 - **Aumento de datos** aumentar la diversidad y la cantidad de las muestras de entrenamiento generando transformaciones aleatorias a las imágenes

existentes, esto se realiza especialmente cuando se tiene un conjunto de datos más pequeño (Google , s.f.)

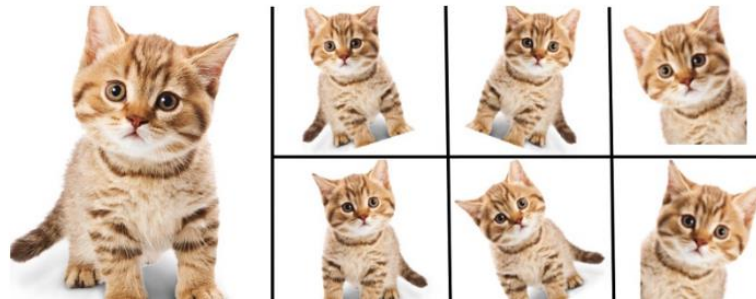


Ilustración 49. Ejemplo de aumento de datos a partir de una imagen (Raj, s.f)

- **Drop Out (abandono):** esta técnica elimina aleatoriamente unidades de las capas (neuronas), permitiendo que no se le asigne mayor peso a unas características que otras, es decir, se aprenda unas características mejores que otras²¹ (Google , s.f.).

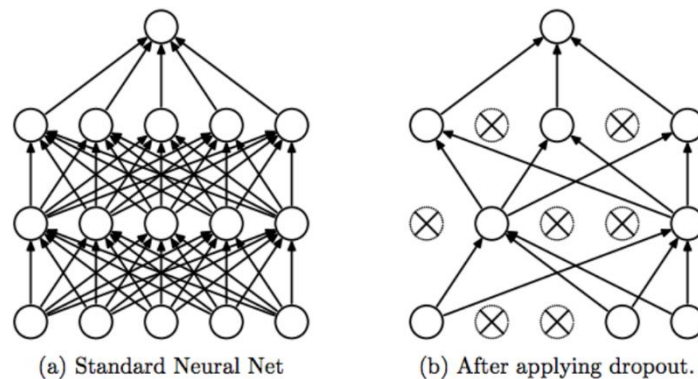


Ilustración 50. Ejemplo de Dropout (Budhiraja, 2016)

- **Detección temprana:** permite que la red no se entrene con tantas iteraciones y así sobre ajustar los valores de los pesos (Ng, s.f.)
- **Capas Densas (Dense):** Al final de la red es necesario tener las capas completamente conectadas, en donde cada nodo de las últimas capas estén conectadas a cada nodo de las capas siguientes, esto se debe realizar varias capas antes de la capa de salida. Adicionalmente la capa final debe tener una función de activación que genere un valor de probabilidad, entre 0 y 1, para cada una de las etiquetas de clasificación que intenta predecir. A esta función se le conoce como softmax (Google , s.f.)

- **Optimizador (Adam):** El optimizador acelera el proceso del descenso del gradiente, permitiendo que llegue más rápido al mínimo sin tantas iteraciones, uno de los optimizadores más utilizados es el algoritmo de Adam, el cual combina otros algoritmos como AdaGrad (algoritmo del gradiente adaptativo) y RMSprop (propagación cuadrada media de la razón). El algoritmo calcula el promedio móvil exponencial del gradiente y del gradiente cuadrado, adicional a esto se utiliza una tasa de aprendizaje en decaída, el cual disminuye en cada época, permitiendo que la tasa de aprendizaje tome valores grandes al principio y que vaya convergiendo con pasos más pequeños (Brownlee, 2017).
- **Pérdida:** El valor que indica que tan incorrecta fue la predicción del modelo, entre menor sea la pérdida mejor es la predicción. (Google, s.f.)

El entrenamiento busca encontrar la curva que mejor se adapte a los datos como se puede ver a continuación.

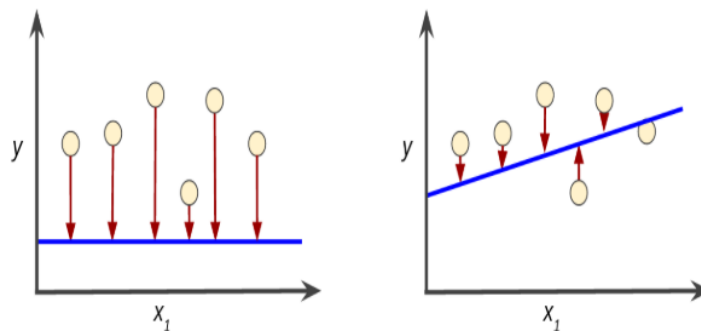


Ilustración 51. Pérdida Alta / Pérdida baja (Google, s.f.)

- **Descenso del gradiente:** Técnica para minimizar la pérdida mediante cálculos de gradientes de pérdida con respecto a los parámetros del modelo (pesos y sesgo) (Google, s.f.).
- **Tasa de aprendizaje:** Valor que se usa para entrenar un modelo a través del descenso de gradientes, buscando el mínimo de la pérdida. (Google, s.f.)
- **Paso de gradiente:** Multiplicación entre la tasa de aprendizaje y el algoritmo del descenso de gradiente. Si el paso es muy pequeño el aprendizaje puede tomar más tiempo para llegar al mínimo, y si el paso es muy grande, es posible que diverja y no alcance a llegar al mínimo. (Google, s.f.)

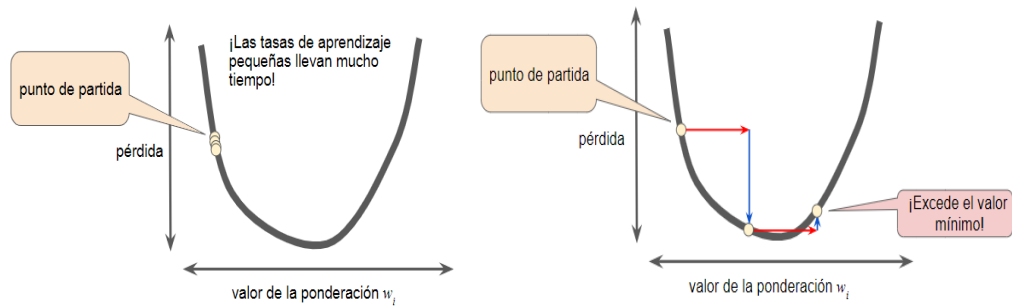


Ilustración 52. Pasos muy pequeños / pasos muy grandes (Google, s.f.)

- **Época:** Recorrido de entrenamiento completo por todo el conjunto de datos, las iteraciones de todo el tamaño del lote (Google, s.f.).
- **Hiperparámetros:** Son valores que permiten controlar los parámetros de la red (pesos y el sesgo); estos hiperparámetros son: número de épocas, número de capas ocultas, número de unidades o neuronas ocultas por capa, la función y activación, valor de la tasa de aprendizaje, el optimizador a utilizar y sus variables.

2. METODOLOGIA

En el sistema de reconocimiento de billetes para personas con discapacidad visual mediante el uso de visión artificial, se usará la metodología Ulrich, la cual consiste en modelos de fases donde estructuran el proceso de diseño en etapa y la secuencia recomendada para llevarlos a cabo. Estos, son modelos de fases orientados al desarrollo del producto (Blanco Romero, 2015)

2.1 DISEÑO CONCEPTUAL

2.1.1 Entrevista a fundaciones

El día 17 de febrero se realizaron una serie de entrevistas a diferentes personas de la Taller Hogar Ángel de Luz ubicada en la ciudad de Medellín; estas entrevistas se realizaron con el fin de conocer las opiniones con respecto a un dispositivo nuevo que se plantea efectuar, y las diferentes necesidades que el dispositivo podría suplir. Con base a esta entrevista se obtuvo una lista de necesidades como se puede ver a continuación, y se definió que los objetos más importantes que se deben analizar son los billetes.

2.2 ESPECIFICACIONES OBJETIVO

2.2.1 Lista de necesidades

En primer lugar se debe aclarar el problema. Para esto, es necesario conocer las necesidades del cliente objetivo y asignarles un valor de importancia que oscila entre los valores del 1 al 5, siendo 5 la mayor importancia y 1 una importancia baja, esta importancia fue dada por cada persona de la fundación, aclarando cuál era su mayor objetivo en el proyecto.

Tabla 3. Lista de necesidades

Número	Necesidad	Importancia
1	El sistema reconoce billetes	5
2	El sistema es cómodo	3
3	El sistema es fácil de usar	4
4	El sistema posee pocos cables	2
5	El sistema es liviano	3
6	El sistema es económico	4
7	El sistema posee bajo consumo de energía	4
8	El sistema responde rápido	4

2.2.2 Lista de Medidas iniciales

Posterior a esto se procede a determinar la unidad de medida correspondiente a cada necesidad, por ejemplo, si la necesidad es que el sistema sea liviano, su unidad de medida serán gramos como se puede ver en la sección, a esto también se le asigna también un valor estimado.

Tabla 4. Lista de medidas

Número de necesidad	Medidas	Importancia	Valor estimado	Unidades
1	Sensibilidad y especificidad	5	5	% error
2,3,4	Nivel de satisfacción	3,4,2	01-may	NA
5	Peso	3	150	G
6	Precio	4	1'000.000	\$
7	consumo de energía	4	3000	mAh
9	Velocidad de respuesta	4	500	ms

2.2.3 Lista de medidas de la competencia

Además de estimar unos valores de medida con sus respectivas unidades, se necesita hacer una comparación entre el dispositivo que se plantea realizar y un dispositivo que ya existe en el mercado, esto con el fin de calcular valores y plantear rango de valores que se desea obtener como se puede ver en la siguiente tabla.

Tabla 5. Lista de medidas de la competencia

Número de necesidad	Medidas	Importancia	Valor estimado	Unidades
1	Sensibilidad y especificidad	5	5	% error
2,3,4	Nivel de satisfacción	3,4,2	01-may	NA
5	Peso	3	1	Kg
6	Precio	4	4'000.000	\$
7	consumo de energía	4	1500	mAh
9	Velocidad de respuesta	4	0.5	ns

2.2.3.1 Productos similares

Uno de los mayores ejemplos de productos similares dispositivo es "Orcam MyEye", el cual es capaz de leer texto, reconocer objetos y rostros previamente almacenados transmitiéndole al usuario un mensaje de voz por medio de un auricular, y todo esto adaptándose a cualquier gafas sin necesidad de comprar gafas especiales y estrambóticas (Essilor, 2016).



Ilustración 53. Orcam MyEye tomado de (Quantumr LV, 2016)

Por otra parte, “EYESYNTH” es un dispositivo más avanzado que el anterior, ya que registra el espacio en tres dimensiones y por medio de un microprocesador convierte la información en un audio comprensible para el invidente el cual ellos llaman “experiencia de sentido aumentado”, este no usa palabras, sino que utiliza un nuevo lenguaje que el cerebro puede entender, el sonido que se trasmite lo hace a través de los huesos de la cabeza sin necesidad de usar auriculares (Eyesynth, s.f.)



Ilustración 54. Eyesynth tomado de (Eyesynth, s.f.)

2.2.4 Lista de medidas ideales y aceptables

Una vez hecho esto, se debe comprar las medidas asignadas a cada necesidad con las de otros dispositivos similares para obtener parámetros estimados entre los que debería estar el dispositivo.

Tabla 6. Lista de medidas ideales y aceptables

Número de necesidad	Medidas	Importancia	Valor ideal	Valor aceptable	Unidades
1	Sensibilidad y especificidad	5	2	10	% error
2,3,4	Nivel de satisfacción	3,4,2	5	>2,<5	NA
5	Peso	3	150	250	g
6	Precio	4	800	2'000000	\$
7	consumo de energía	4	1500	3000	mAh
9	Velocidad de respuesta	4	20	500	ms

2.2.5 Matriz de necesidades y medidas

Para concluir se construyó una matriz de Necesidades vs medidas y se le asignándole a cada necesidad una medida

Tabla 7. Matriz de necesidades/ medidas



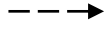
Matriz de necesidades /medidas		1	2	3	4	5	6
		Sensibilidad y especificidad	Nivel de satisfacción	Peso	Precio	Consumo de energía	Velocidad de respuesta
1	El sistema reconoce objetos						
2	El sistema es cómodo						
3	El sistema es fácil de usar						
4	El sistema posee pocos cables						
5	El sistema es liviano						
6	El sistema es económico						

7	El sistema posee bajo consumo de energía						
8	El sistema responde rápido						

2.3 DISEÑO CONCEPTUAL

2.3.1 Aclaración del problema

Para aclarar el problema, se debe verificar tanto las entradas como las salidas del sistema mediante el uso de la caja negra y del diagrama de funciones. Esto se hace con el fin de comprender el funcionamiento del mismo. Cabe anotar que para este procedimiento se utilizan convenciones como las que se muestran a continuación:

-  Las flechas delgadas continuas representan los flujos de energía, como la energía que entra al sistema, o la potencial y cinética que poseen los dispositivos.
-  Las flechas gruesas y continuas representan el material requerido para la función del dispositivo.
-  Las flechas delgadas y discontinuas representan las señales como la salida de sensores o el sistema de alertas para seguridad.

2.3.1.1 Caja negra



Ilustración 55. Caja negra

2.3.1.2 Diagrama de funciones

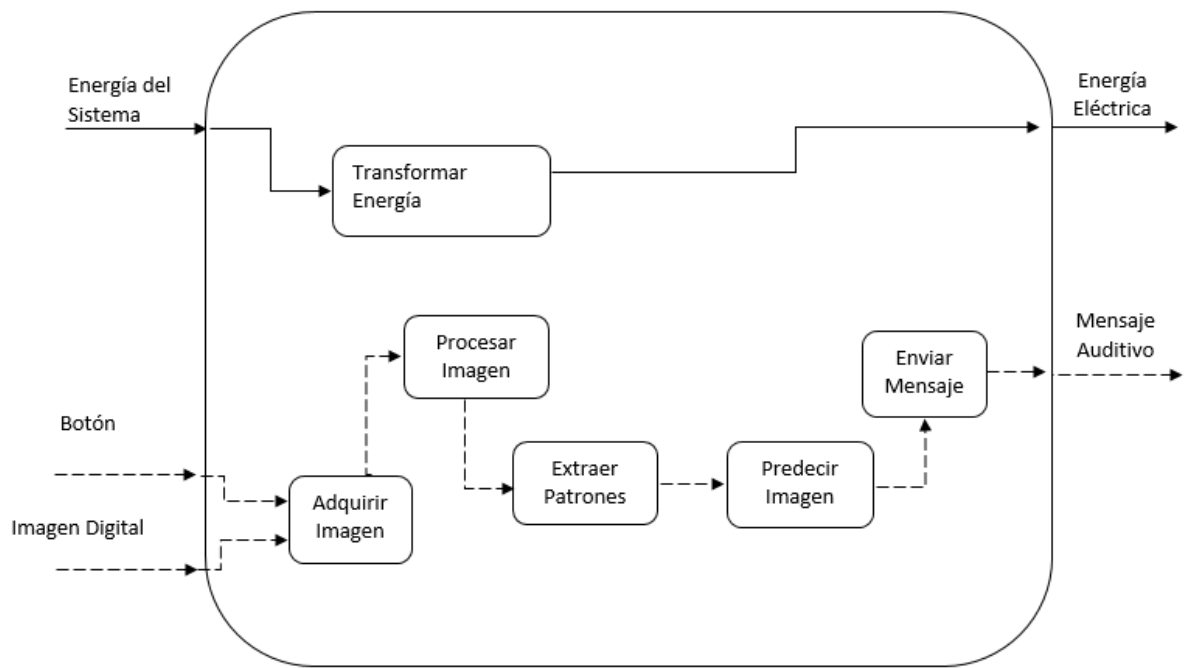


Ilustración 56. Diagrama de funciones

Una vez se tiene el diagrama de funciones, se debe verificar, internamente, con qué tipo de elementos se puede llevar a cabo cada una de las funciones del dispositivo. Estos elementos, junto con su descripción, se deben consignar en el registro de soluciones para verificar que no existan dos o más elementos que puedan llevar a cabo la misma función.

2.3.2 Registro de soluciones

Luego de verificar la aclaración del problema se debe acomodar en la matriz morfológica las posibles soluciones que cumplen cada una de las funciones, para obtener así un conjunto de soluciones de las cuales se puede elegir cuál es la más viable mediante un análisis de pros y contras de cada una de ellas. Además será necesario ubicar las soluciones en una matriz de ponderación para verificar, mediante la suma de sus bondades y desventajas, cuál de ellas se debe desarrollar.

2.3.2.1 Matriz morfológica

La matriz morfológica ubicada en el Anexo 1, se realiza con el objetivo de escoger de la mejor propuesta de componentes y/o materiales para llevar a cabo el dispositivo, con respecto las funciones establecidas en el diagrama anterior.

2.3.2.2 Definición de conceptos

Tabla 8. Definición de conceptos

Concepto	Definición
A	La estructura seleccionada para el concepto A se basa en capturar la imagen utilizando una mini webcam, la cual será pre procesada por una Raspberry que tendrá como fuente de energía una batería de litio. El software que utilizará para extraer características y realizar predicciones es Python, para finalmente enviar el resultado por medio de unos audífonos alámbricos.
B	Para el concepto B se planteó como solución utilizar gafas con cámara para adquirir la imagen, que será pre procesada en un computador, el cual tendrá una batería de níquel hidruro. Este concepto permitirá extraer características y predecir el tipo de objeto mediante el software Matlab, proporcionando resultados a auriculares con radio frecuencia.
C	Para el concepto C, se utilizará una webcam para adquirir la imagen que será pre procesada en un sistema embebido llamado Odroid, el cual su fuente de energía será una batería Lipo. El sistema realizará extracción de características y predicción de objetos mediante el software C++, enviando un resultado final por medio de unos auriculares bluetooth.

2.3.2.3 Selección de concepto

La selección de conceptos fueron evaluados en la siguiente tabla tomando como criterio de selección la lista de necesidades, la escala de evaluación se realiza mediante selecciones discretas:

- “+”: el concepto cumple con el criterio
- “0”: el concepto es indeterminado para el criterio
- “-”: el concepto no cumple con el criterio

Para realizar la ponderación de los criterios con “+” aportan importancia al concepto (suman), los criterios con “0”, son indiferentes al concepto y los criterios con “-”, quitan importancia al proceso (restar)

Tabla 9. Selección de concepto

Criterio de selección	Conceptos		
	A	B	C
Menor % error	+	+	+

Comodidad	0	-	-
Fácil de usar	+	+	+
Bajo peso (liviano)	+	-	+
Económico	+	-	0
Bajo consumo de energía	+	-	0
Velocidad de respuesta alta	0	+	+
Suma +	5	3	4
Suma 0	2	0	2
Suma -	0	4	1
Evaluación Total	5	-1	3
Escala	1	3	2

2.3.3 Diseño de detalle

El diseño de detalle consta de dos partes, el diseño de hardware y el diseño de software.

2.3.3.1 Diseño de hardware

El diseño de hardware comprende los diferentes mecanismos que permiten un adecuado funcionamiento del sistema (fuente de alimentación, cámara y sistema embebido). Para integrar todas las partes se consideraron 2 prototipos experimentales, uno basado en un sistema de iluminación controlado totalmente, y otro basado en un sistema abierto controlado parcialmente mediante un patrón característico en la escena. Para ambos diseños se consideran las siguientes etapas:

2.3.3.1.1 Fuente de alimentación

Dado el concepto A, para transformar energía se eligió batería de litio bajo la siguiente tabla general de consumo encontrada en los elementos que conforman el sistema:

Tabla 10. Características de fuente de alimentación

Elemento	Características
Iluminación led	120 - 180 mA/h
Cámara	100 – 200 mA/h
Procesador	2000 - 2500 mA/h

Total	2220 – 2880 mA
--------------	-----------------------

La batería Litio seleccionada es la referencia 18650 integrada en un sunfounder con entrada de 7V, 4000 mA/h y salida de 5V, 2A (Amazon, 2016), con la cual estimando un uso continuo se espera una duración de 4 horas.

2.3.3.1.2 Cámara

Dado el concepto A, la cámara se seleccionó bajo las características mínimas requeridas para el sistema de tal forma que el tiempo de procesamiento fuera mínimo, además de cumplir con su fácil integración a las gafas en peso y tamaño.

Las características mínimas se presentan en la siguiente tabla:

Tabla 11. Características de la cámara

Característica	Descripción
Resolución	1600 x 1200, 1280 x 1024, 1280 x 720, 640 x 480
Ángulo de visión	54°
Longitud focal	6 mm
Captura de imágenes	30 FPS
Peso	20 g
Tamaño	7 x 32 mm (D x L)
Tamaño sensor	0.28 mm

2.3.3.1.3 Sistema Embebido

Seleccionó la raspberry pi2 modelo B debido a que sus características permiten el procesamiento de imagen necesario para el sistema, además de cumplir con el concepto A en peso, disponibilidad y portabilidad.

Además, luego de tener el proyecto terminado se podría migrar está a una raspberry mini para mayor comodidad del usuario. En la siguiente tabla se muestran las especificaciones del producto (RASPBERRY PI FOUNDATION, 2015):

Tabla 12. Características sistema embebido

Característica	Descripción
RAM	1GB
Procesador	A 900MHz quad-core ARM Cortex-A7 CPU
Ethernet	Si
USB	Si
HDMI	Si
GPIO pins	Si

2.3.3.1.4 Prototipo 1 con sistema de iluminación controlado totalmente

Para tener un sistema de iluminación controlado se debe ubicar la cámara dentro de una caja o gabinete de tal forma que sea posible tener un buen enfoque del billete y un control estable de la iluminación.

La ubicación de la altura de la cámara respecto al billete, conocida como distancia de trabajo, se deberá calcular a partir de la siguiente fórmula:

$$HFOV = 2 * WD * \tan\left(\frac{AFOV}{2}\right),$$

Donde HFOV es la profundidad de campo horizontal, WD es la distancia de trabajo y AFOV es el ángulo de visión. Al despejar WD se obtiene la siguiente fórmula:

$$WD = \frac{HFOV}{2 * \tan\left(\frac{AFOV}{2}\right)}$$

Conservando un offset de 50 mm a cada lado del billete se calcula un distancia de trabajo mínima de $WD = 248$ mm y con este valor se diseña en material MDF con un grosor de 3 mm de espesor el cual cuenta con dimensiones de 282 mm de alto (considerando los 32 mm de longitud de la cámara), 158 mm de ancho y 76 mm de profundidad y una ranura de 10mm de altura sobre la base inferior para ingresar y extraer el billete fácilmente, estas dimensiones se pueden ver en la siguiente figura.

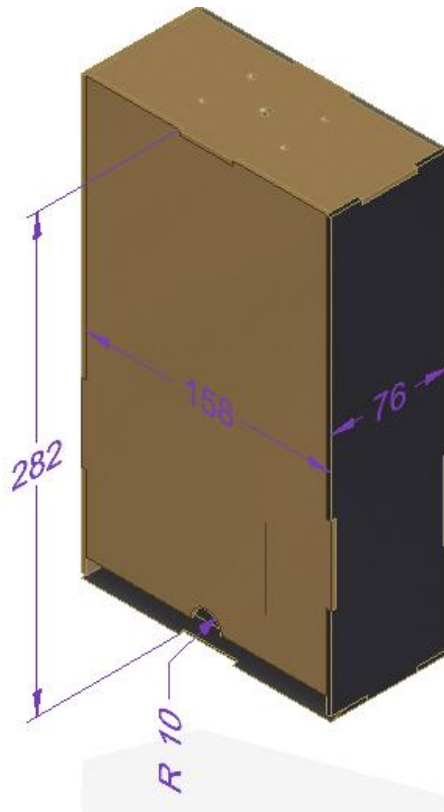


Ilustración 57. Dimensiones del diseño físico del sistema controlado

La caja al interior está pintada completamente de negro para generar un ambiente controlado como se puede ver en la figura 41. La iluminación se adecuó con 4 LED's de color blanco ubicados alrededor de la cámara como se ve en la siguiente figura. Esta iluminación tiene como nombre iluminación frontal en donde la luz se sitúa en la misma dirección que la cámara

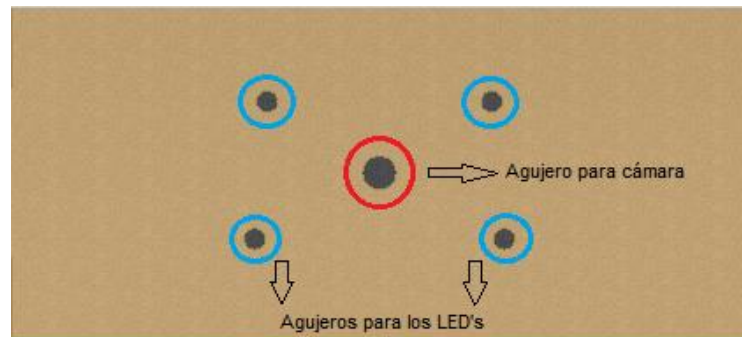


Ilustración 58. Vista superior del sistema



Ilustración 59.vista exterior del sistema controlado

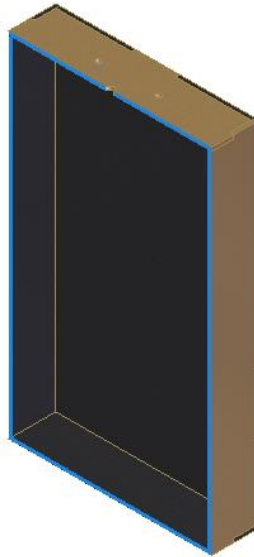


Ilustración 60. Vista interna del sistema controlado

2.3.3.1.2 Prototipo 2 con sistema de iluminación controlado parcialmente

Para el segundo prototipo se consideró en primer lugar un soporte para ubicar la cámara junto a las gafas, y en segundo lugar un guante con un patrón característico.

Diseño soporte Cámara

- Soporte para la mini webcam que es de forma cilíndrica y debe adherirse a la pata lateral derecha las gafas.

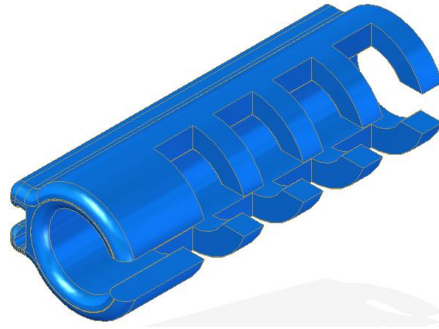


Ilustración 61. Soporte de cámara para gafas

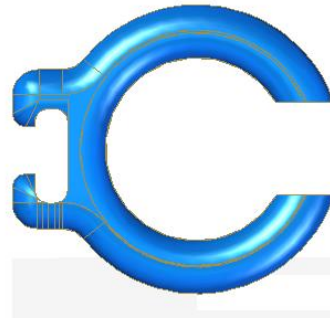


Ilustración 62. Vista frontal soporte cámara

El soporte de la cámara se diseñó para ser impresa en una impresora 3D en material ABS, con dimensiones de 6.9 mm de diámetro (3,45 de radio), con una longitud de 32 mm y para acoplar a las gafas se utilizó una altura 3.56 mm, como se puede ver en la siguiente figura:

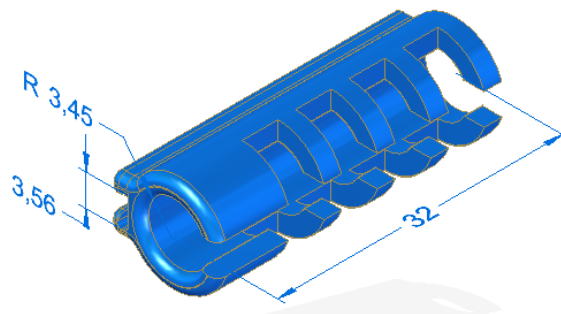


Ilustración 63. Dimensiones soporte de cámara

Diseño guante patrón

La selección del guante se establece a partir de un diseño experimental en el que se seleccionan los siguientes atributos:

- Color rojo en todo el guante sin ranura para ubicar el billete
- Color rojo en todo el guante con ranura para ubicar el billete
- Color negro con borde verde con ranura para ubicar el billete

El primer experimento se basa en sujetar el billete de manera natural de tal forma que este quede sobre el guante de color rojo como se muestra en la figura 64:



Ilustración 64. Guante rojo

El segundo experimento consiste en ubicar el billete en una ranura interna del guante, quedando este lo más centrado posible respecto al área del guante como se puede observar en la figura 65:



Ilustración 65. Guante diseño 2/ Parte palmeada Bolsillo enmarcado

El tercer diseño es una un guante con bolsillo similar al anterior, sin embargo en este se considera un contraste de colores entre el billete y el guante



Ilustración 66. Guante negro de algodón con bordes verdes

2.3.3.2 Diseño de software

2.3.3.2.1 Diseño de Software para prototipo 1

Para la clasificación de los billetes en la estructura del prototipo 1 se cuenta con las siguientes etapas:

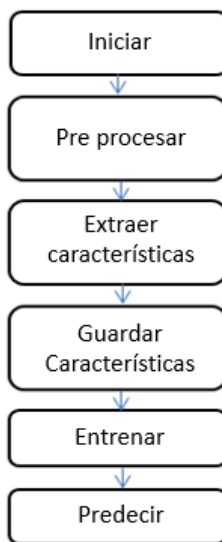


Ilustración 67. Estructura del código

Captura de imágenes

La captura de imágenes es la etapa inicial para comenzar con la implementación de un proyecto de visión artificial, esta representa la cantidad y el contenido de la información que se le proporciona más adelante al modelo de predicción. Para ello se debe tener en cuenta

la cantidad de objetos que se quiere clasificar, el entorno en dónde se pretende trabajar, y la cantidad de imágenes que se debe tomar por objeto.

- Los objetos que se quieren clasificar son los billetes colombianos nuevos y viejos.
- La cantidad de imágenes que se debe capturar tiene una relación con el ítem anterior, entre mayor sea la cantidad de objetos a clasificar mayor debe ser la cantidad de capturas por objeto y además debe ser uniforme.

Luego de tener estos ítems en cuenta, Las imágenes se deben capturar de manera diferenciada, por ejemplo, capturar imágenes en diferentes variables como:

- Tipos de iluminación: con poca, media o mucha luz que capta la cámara
- Tamaños del objeto: ya sea por la distancia de la cámara al objeto o por que los objetos tienen diferentes tamaños
- Formas: en la inclinación, perspectiva o relación de aspecto
- Entornos: ubicar el objeto en diferentes escenarios.

Pre procesamiento

El pre procesamiento es la etapa siguiente a la captura de imágenes (inicio de la programación), la cual realiza modificaciones a la imagen para segmentar el objeto de interés.

El diseño del pre procesamiento se puede ver en el Anexo 2, el cual muestra un diagrama de flujo para realizar posteriormente el algoritmo.

Extracción de características

Una vez obtenida la región segmentada, se extraen características, las que diferencian a los objetos en color, forma, tamaño, textura o contenido; estas son muy relevantes a la hora de trabajar con reconocimiento de imágenes, puesto que son estos valores los que aprende la red.

Cuando se requiere clasificar objetos es importante conocer cómo se diferencian entre ellos, pues en algunos objetos para hallar la mejor o las mejores características no resulta tan evidente, en muchos casos es imperceptible, por eso se plantean varias formas para hallar las mejores características:

- **Promedio del valor de pixel en gris por filas:** se considera la siguiente imagen:

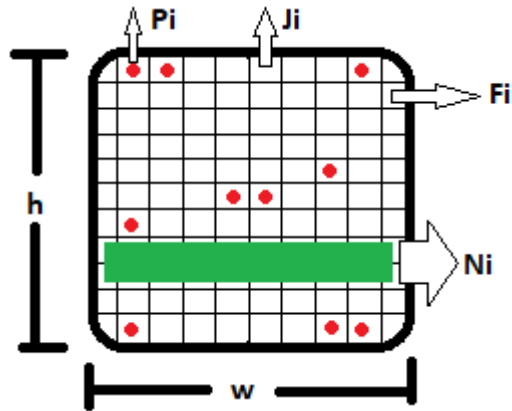


Ilustración 68. Promedio valor de pixel en gris por filas

Los valores de la imagen corresponden a:

- h es el alto de la imagen
- w es el ancho de la imagen
- F_i es la fila en la posición i
- J_i es la columna en la posición j
- P_i es el valor del pixel en la posición (F_i, J_i) de la imagen
- N_i es el # de agrupación de filas que determinan el número de características obtenidas en la imagen

El promedio del valor de pixel en gris estará dado por:

$$\bar{m}_i = \frac{1}{N_i * w} \sum_{j=0}^{w-1} P_i[N_i]$$

Promedio del valor de pixel en gris por bloques de filas y columnas: Similar al anterior consiste en recorrer la imagen en gris promediando los valores de pixeles por bloques B_i calculando la cantidad de características de la siguiente manera:

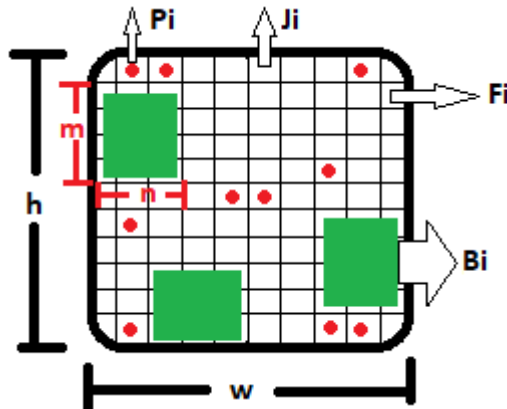


Ilustración 69. Promedio del valor de pixel en gris por bloques de filas y columnas

$$\bar{m}_i = \frac{1}{m * n} \sum_{j=0}^{n-1} Pi[Bi]$$

- **Promedio de histograma en RGB:** Consiste en separar en imágenes diferentes los canales RGB y calcular el vector de histogramas de cada una de esas imágenes, los valores de cada uno de los vectores, se promedia la cantidad deseada, guardando cada promedio como una característica

*Cantidad de características = Porcentaje de valores a promediar * ancho del vector * cantidad de canales*

- **Promedio de histograma en HUE:** Consiste en separar en imágenes diferentes los canales de HSV, calculando el vector histograma de solo el canal de color (HUE), los valores del vector se promedian una cantidad deseada, guardando cada promedio como una característica

*Cantidad de características = Porcentaje de valores a promediar * ancho del vector*

- **Local binary patterns (LBP):** Esta es una característica de textura, invariante a la rotación, que analiza los vecinos alrededor dependiendo del pixel en el que este parado en el instante

Consiste en a partir de la imagen en gris se extrae una porción de la imagen y se analiza los valores alrededor del pixel central como se puede ver en la figura

9	3	8
3	7	9
8	8	2

Ilustración 70. Vecinos del valor central

Cada valor del pixel tiene una posición correspondiente al analizar los valores del local binary patterns

$$\left\{ \begin{array}{l} 1 \text{ si valor del vecino} \geq \text{Valor central} \\ 0 \text{ en otro caso} \end{array} \right\}$$

Se visualiza de la siguiente manera:

1	0	1
0	7	1
1	1	0

Ilustración 71. Patrones local binary patterns

Estos valores se despliegan para hallar en código

1	0	1	1	0	1	1	0
---	---	---	---	---	---	---	---

Ilustración 72. Posición de local binary patterns

Se realiza el siguiente cálculo convirtiendo de binario a decimal.

$$2^7 + 2^5 + 2^4 + 2^2 + 2^1 = 182$$

El valor en decimal se le asigna al valor central de la porción de imagen que se analizó, esto repite hasta llegar a llenar toda la imagen

	182	

Ilustración 73. Valor central con nuevo código

Convirtiendo toda la imagen a valores del local binary patterns

Existen varias variaciones para lo local binary patterns para generar robustez, como reducir los cambios monotónicos del nivel de gris, asignando un umbral en vez de igualarlo a cero, como se puede ver en la imagen

$$\begin{aligned}
 & \mathbf{1} \text{ si } \text{valor del vecino} - \text{Valor central} \geq T \\
 & \mathbf{0} \text{ en otro caso}
 \end{aligned}$$

A partir de la imagen LBP se hallan los valores del histograma como características, sin embargo cuando se hallan los códigos LBP hay patrones que no tienen probabilidad de darse

Entonces se creó el método uniforme, En donde la U representa el número de transiciones que hay para pasar del 0 al 1 o del 1 al 0 en representación binaria, por ejemplo:

- Cuando U= 0 significa que no hay cambios de transiciones

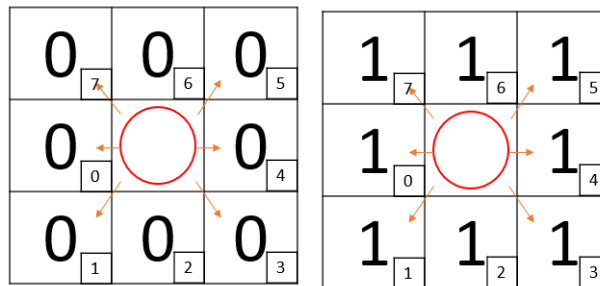


Ilustración 74. Ejemplos de patrones cuando U=0

- Cuando U = 2 significa que hay 2 cambios

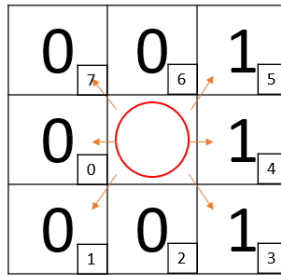


Ilustración 75. Ejemplo de patrones cuando $U=2$

- Cuando $U=4$ significa que hay 4 cambios

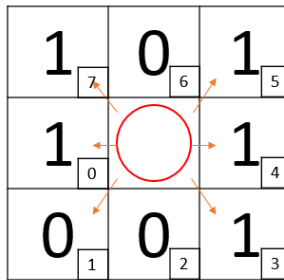


Ilustración 76. Ejemplo de patrones cuando $U=4$

Los LBP son uniformes cuando:

- Los patrones con $U=0$ y $U=2$ se les asigna un código nuevo
- Los patrones diferentes a estos dos se les asigna el mismo código

Entonces los patrones se reducen a 59 patrones, ya que 29 valores corresponden a las posibles combinaciones de $U=0$ y otras 29 corresponden a combinaciones de $U=2$ y 1 corresponde al código que se le asignó al resto.

Para cada código de LBP representa un tipo de característica en la imagen como se puede ver en la figura.

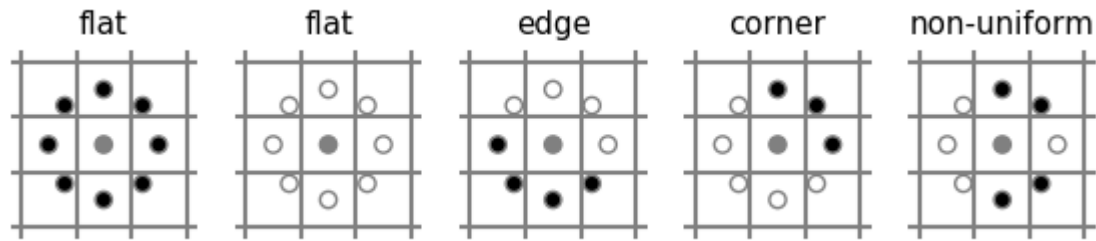


Ilustración 77. Patrones con sus respectivas tipos de características

Para finalizar se calcula histogramas en LBP y dependiendo del radio que se elija (distancia en pixel del valor central hacia los vecinos), el vector tendrá cierto tamaño. Cada valor del histograma se guarda en un vector como una característica

Después de realizar la extracción de características, se procede a guardarlas, como se puede ver a continuación.

Almacenamiento de características

Para explicar cómo se guardaron los datos se deben aclarar los siguientes términos:

En este documento se menciona el término “clase” cuando hacen referencia a algún objeto en específico.

Cuando se extraen las características de cada billete y se ingresan los valores en un vector de características, ese vector se guarda en un archivo que se llena con cada imagen que es pre procesado para que al final continúe al siguiente proceso. El archivo debe tener:

- Las características de cada billete (representa las entrada)
- La clase de cada una de las imágenes (representa la etiqueta)

X_n : características
 Y_m : etiquetas

El archivo generado en la primera columna, almacena la clase y en las otras columnas las características correspondientes a cada muestra, como se puede ver en la siguiente tabla.

Tabla 13. Distribución de clases y características

Clase / Característica	X_0	X_1	X_2	. . .	X_n
Y_0	c_{00}	c_{01}	c_{20}	. . .	c_{0n}
Y_1	c_{01}	c_{11}	c_{12}	. . .	c_{1n}
Y_2	c_{21}	c_{21}	c_{22}	. . .	c_{2n}
.
.
.
Y_m	c_{m1}	c_{m1}	c_{m2}	c_{11}	c_{mn}

Existen diferentes maneras de almacenar los datos, una de ellas es mediante archivos Excel, puesto que se pueden visualizar los valores más fácilmente, otra manera en la que se pueden guardar son archivos “.txt” o “.pickle”, que se pueden abrir con un bloc de notas para visualizarse.

Entrenamiento

Entrenamiento es el proceso a través del cual se le enseña a la red neural a realizar una clasificación de imágenes.

Los valores de características y etiquetas deben estar guardados en un vector de etiquetas y una matriz de características, se procede a normalizar los datos, estandarizando las características de forma que se elimina la media y se escalona la varianza, estas dos cosas las hace de manera independiente.

Luego de realizar una normalización, se procede a dividir el conjunto de los datos en subconjuntos de pruebas y de validación, esta división se realiza con el fin de conocer la capacidad de generalización del clasificador, es decir, conocer cómo se comporta el modelo con datos que no corresponden al entrenamiento y evitar el sobre ajuste de la red.

Usualmente se usa un subconjunto mayor para el entrenamiento y una cantidad diversa de datos para la validación, el conjunto de muestras se puede dividir de las siguientes formas:

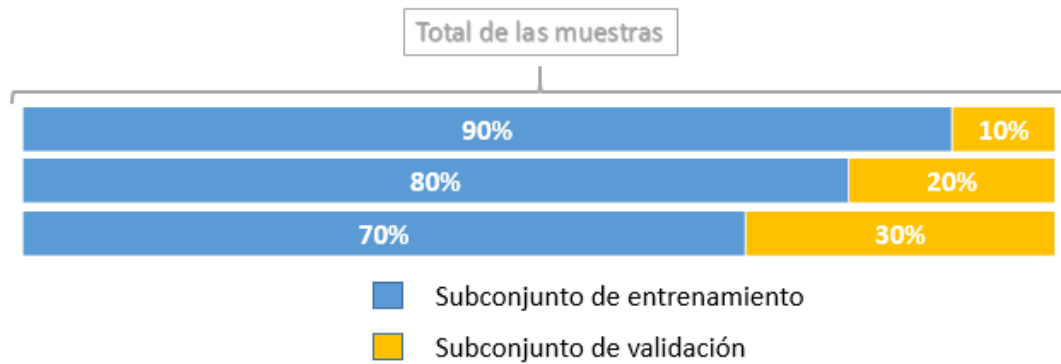


Ilustración 78. Diferentes maneras de separar el conjunto de entrenamiento y el conjunto de validación

Estas son las muestras más comunes que pueden haber, entre otras, lo que se espera es que el conjunto de entrenamiento sea mayor al de validación; entre más datos se tenga para entrenar más confiable es el modelo.

Después de dividir el conjunto de datos, se procede a entrenar los datos, para hallar el mejor modelo, utilizando diferentes hiperparámetros, los cuales son los valores configurables de la red

Luego se realiza una matriz de confusión con los datos de validación para comprobar qué tanto se equivoca la red y en cuales clases con mayor frecuencia.

La matriz de confusión se utiliza para conocer la calidad de la salida de un clasificador, los elementos diagonales representan el número de puntos para los que la predicción es igual a los valores reales, mientras que los elementos fuera de la diagonal son aquellos que están mal etiquetados por el clasificador, entre mayor sean los valores de la diagonal, mejor es la predicción.

Tabla 14. Matriz de confusión

		Valores predijo	
		Clase 1	No – Clase 1
Valores reales	Clase 1	Reales Positivos (RP)	Falsos Negativos (FN)

A continuación, se presenta un ejemplo con valores de la matriz de confusión de diferentes objetos:

Tabla 15. Ejemplo de matriz de confusión para algunos billetes

		Valores predijo		
		Persona	Carro	Moto
Valores reales	Persona	5	2	0
	Carro	3	4	2
	Moto	0	1	7

Luego se analiza la tabla de confusión para todas las clases, Ejemplo: (“Persona”)

Tabla 16. Análisis de la matriz de confusión anterior para una clase

		Valores predijo	
		Persona	No Persona
Valores reales	Persona	RP: 5	FN: 2
	No - Persona	FP: 3	RN: 14

A partir de esta matriz se halla la precisión y exactitud, la precisión define qué tan cerca está los valores que predijo de una clase, y la exactitud representa que tan cerca está el valor predijo del valor real. Los cálculos se realizan de la siguiente forma tomando en cuenta los valores de la matriz de confusión.

$$Exactitud = \frac{Reales Positivos + Reales Negativos}{Predicciones totales}$$

$$Precision = \frac{5 + 14}{24} = 0.79$$

$$\textit{Precision} = \frac{\textit{Reales Positivos}}{\textit{Reales Positivos} + \textit{Falsos Positivos}}$$

$$\textit{Precision} = \frac{5}{5 + 3} = 0.62$$

Como la precisión y exactitud se miden en porcentaje

$$\textit{Exactitud} = 0.79 * 100 = 79\%$$

$$\textit{Precision} = 0.62 * 100 = 62\%$$

También con estos valores se puede hallar la sensibilidad y especificidad

$$\textit{Sensibilidad} = \frac{\textit{Reales Positivos}}{\textit{Reales Positivos} + \textit{Falsos Negativos}}$$

$$\textit{Precision} = \frac{5}{5 + 2} = 0.71$$

$$\textit{Especificidad} = \frac{\textit{Reales Negativos}}{\textit{Reales Negativos} + \textit{Falsos Positivos}}$$

$$\textit{Precision} = \frac{14}{3 + 14} = 0.82$$

La tasa de reales positivos se define como la sensibilidad y la tasa de falsos positivos es la especificidad.

Uno de los mayores errores que se comenten a la hora de entrenar el modelo es confiar sólo en los valores que de la exactitud de la red con los datos que previamente se entrenó, porque existe la posibilidad de que al partir los datos de manera aleatoria con que se realizó previamente no sea la mejor partición para entrenar la red, y esto genere un sobre ajuste en la red, además de que se reduce la cantidad de muestras que posiblemente habrían mejorado los resultados del modelo. Para evitar esto, se crean mecanismos como es el cross validation o la validación cruzada

El cross validation es un mecanismo divide los datos en conjuntos de entrenamiento y validación, de la siguiente manera:

- Un modelo se entrena usando k-1 conjuntos como datos de entrenamiento

- El modelo resultante se valida en la parte restante de los datos

Esto se puede ver en la siguiente figura, en donde los datos se dividen en conjunto de entrenamiento y validación, se evalúa el modelo con esa partición realiza la acción hasta que el haya pasado por todas las muestras, como se puede ver en la figura.

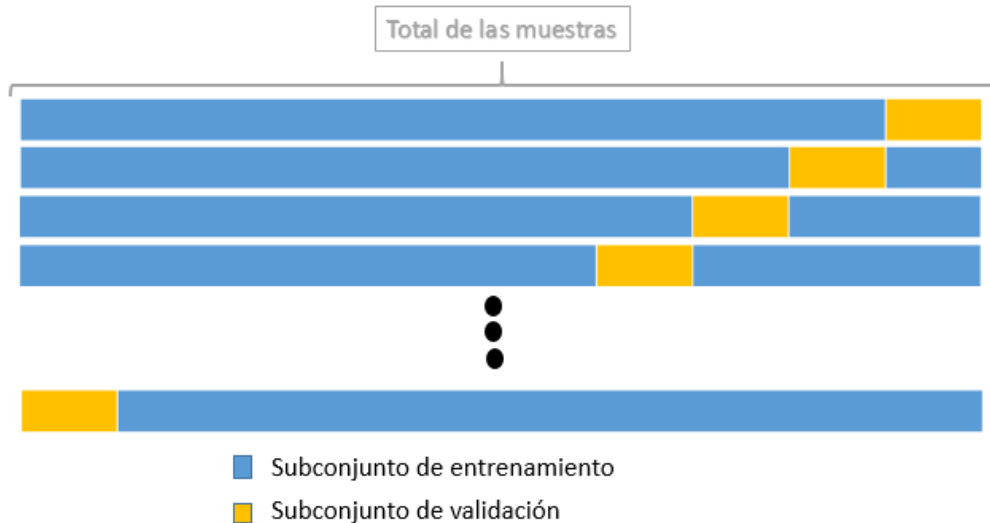


Ilustración 79. Ejemplo validación cruzada

Como resultado final se busca obtener un modelo con que cumpla con requisitos tales como: un valor de exactitud alto presentado en el cross validation, el mínimo de valores fuera de la diagonal de la matriz de confusión y alta capacidad clasificar datos diferentes a los aprendidos.

Luego de realizar diferentes pruebas, se compara y se escoge el mejor modelo, se procede a guardarlo para continuar con el último paso.

Predicción

En el Anexo 3 se puede observar cómo se realiza la predicción y reconocimiento de la imagen.

Inicialmente se cargan los archivos que se guardaron previamente (archivos del modelo) y se activa la cámara, luego se realiza un bucle infinito que permite capturar las imágenes en tiempo real y predecir cuántas veces se quiera. Para ello se debe ingresar el billete, y presionar un botón que le indique al sistema el momento en el que debe comenzar a predecir.

La predicción se realiza se la siguiente manera:

- Se realiza un pre procesamiento de la imagen

- Se extraen las características de la misma manera que se realizó para el entrenamiento
- Se utilizan los modelos previamente entrenados para realizar la predicción, el cual arroja un vector de probabilidades de cada clase.
- Se analiza la posición en el vector con la mayor probabilidad. Si cumple la condición de ser mayor que un porcentaje asignado (en este caso 80%), la predicción se acepta y el proceso continúa, si no la predicción se rechaza informando que no fue posible reconocer el billete, por medio de una voz audible.
- Aceptada la predicción, el programa indica la denominación del billete, por medio de una voz audible.

Finalmente, el programa vuelve a un estado de espera en donde se debe presionar el botón de predicción o el de salir del programa para que realice las tareas respectivas.

2.3.3.2 Diseño de software para prototipo 2

Dado a que se plantearon 3 diseños de guantes, para cada uno de ellos se diseñó un software, algunos presentan la estructura lógica que se utilizó en el prototipo 1.

- **Pre procesamiento**

El diseño del pre procesamiento para los experimentos 1, 2 y 3 se expone en los anexos 4, 5 y 6 respectivamente.

- **Extracción de características, almacenamiento de características y Entrenamiento**

La secuencia lógica se realizó de la misma manera que la del prototipo anterior. Los ítems de extracción, almacenamiento de características y entrenamiento son iguales para los primeros 2 diseños de guante, el tercero presenta una secuencia diferente utilizando otra técnica de clasificación conocida como Deep learning.

La clasificación mediante Deep learning se modela a continuación:

- **Definición del Modelo**

El modelo se diseñó con la siguiente estructura

- Se aplica el filtro convolucional, definiendo el ancho, alto y de una profundidad predefinida (conv2d).
- Se aplica la función de activación (Relu)
- Se normalizan los datos (Batch normalization)

- Se agrupan los máximos valores, buscando las características de mayor valor (Maxpooling)
- Se separan algunas unidades ocultas (Dropout)
- Se repiten los pasos anteriores para formar nuevas capas y volver la red más profunda
- Finalmente se realiza un aplanamiento de los datos (flatten)
- Se unen todas las neuronas de la última capa (Dense)
- Se aplica función de activación (softmax)

El anexo 7 muestra con mayor detalle las entradas y salidas de la red

- **Compilación del modelo**

En la configuración de la compilación del modelo, es necesario asignar la forma en la que se calcular la pérdida dependiendo de las clases que se tengan, adicional a esto es necesario escoger el optimizador.

- **Entrenar el modelo**

Para realizar el entrenamiento se deben separar los datos en subconjuntos de entrenamiento y validación, realizar un aumento de datos al conjunto de entrenamiento para evitar el sobre ajuste y al final del entrenamiento verificar con datos de validación.

Después de realizar los pasos anteriores se procede a configurar la cantidad de épocas con la que se quiere entrenar el modelo, para finalmente realizar dicho entrenamiento, y al hacer varias pruebas cambiando hiperparámetros se elige el mejor modelo y se guarda para utilizarlo en la predicción.

- **Predicción**

La predicción en los tres diseños es de manera similar, en donde se cargan los archivos del modelo, en un bucle infinito se realizan el pre procesado necesario, luego se realiza una captura de la imagen en tiempo real y se predice: Esta predicción arroja un vector con la cantidad de probabilidad que existe para cada clase, se toma la predicción de la mayor probabilidad y si la predicción es mayor a un porcentaje predefinido, esa clase es la denominación, luego se realiza de manera similar las partes que siguen de la predicción.

3. PRESENTACIÓN Y DISCUSIÓN DE RESULTADOS

3.1 PROTOTIPO FUNCIONAL

3.1.1 Construcción prototipo 1

El prototipo funcional del diseño físico se puede ver a continuación:



Ilustración 80. Foto del sistema controlado

A partir de este prototipo se realizaron pruebas con el código para comprobar su funcionamiento, tomando en cuenta los pasos indicados en el presente trabajo de grado.

Inicialmente se realizó una captura de imágenes teniendo en cuenta los factores mencionados en la sección de captura de imágenes, en donde se tomaron al rededor 700 muestras para cada billete, excepto por el billete de \$100.000, el cual queda por fuera de clasificación debido a la escasa circulación y dificultad en adquirir una cantidad suficiente de muestras.

Pre procesamiento

La imagen ingresa con un tamaño de 640X480 pixeles, luego se le aplica un cambio de tamaño para llevarlas todas a un tamaño estándar que consume menos cálculos, luego se convierte a escala de grises y se le aplica un filtro gaussiano.

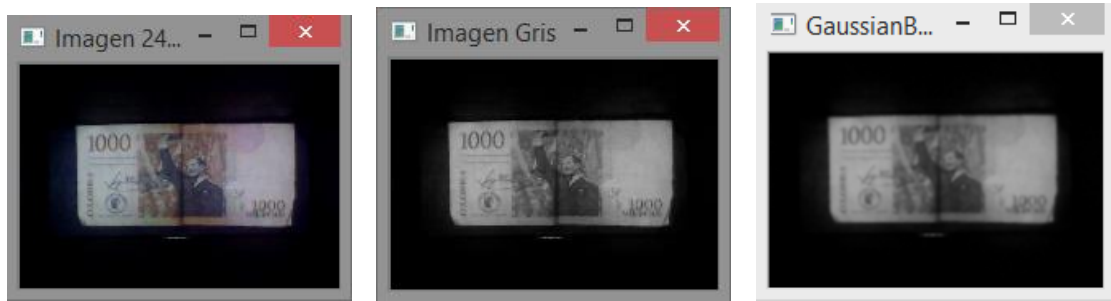


Ilustración 81. Imagen color/ imagen gris/ imagen filtrada

Después se binariza de forma automática, en donde internamente la función obtiene dos umbrales máximos por el cálculo del histograma para luego llevar a imagen a un resultado como el siguiente

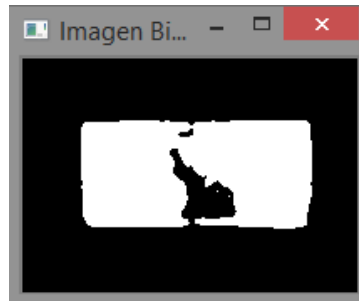


Ilustración 82. Imagen binarizada

Se puede apreciar en la imagen que hay partes que se cortan y no completa el cuadrado del billete, esto se debe a factores como la iluminación o la Umbralización que se calculó. Para resolver este problema se decidió dilatar la imagen con un kernel cuadrado para cerrar los huecos que tiene la imagen anterior y poder completar el contorno y así poder rellenar el contorno de mayor área para eliminar el ruido interno

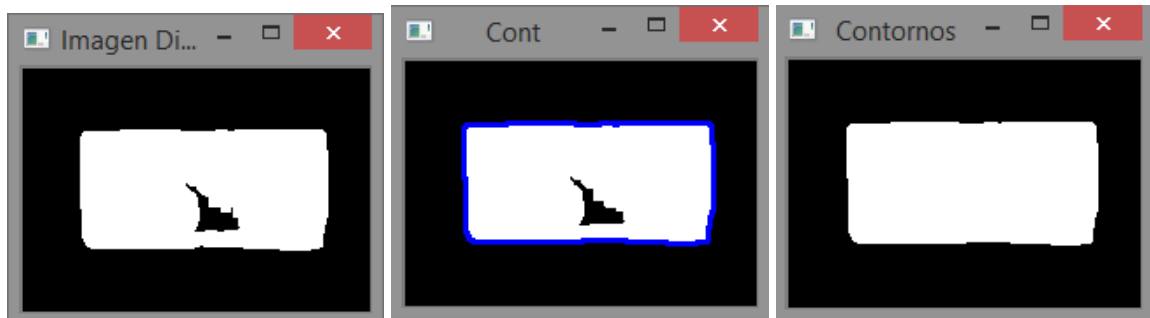


Ilustración 83. Imagen dilatada / contorno de imagen/ imagen rellena

Luego se le aplica una aproximación del contorno buscando la forma de un rectángulo el cual contiene 4 vértices. Al obtener un resultado similar a la forma geométrica del billete se procede a dibujar la figura como se observa a continuación:

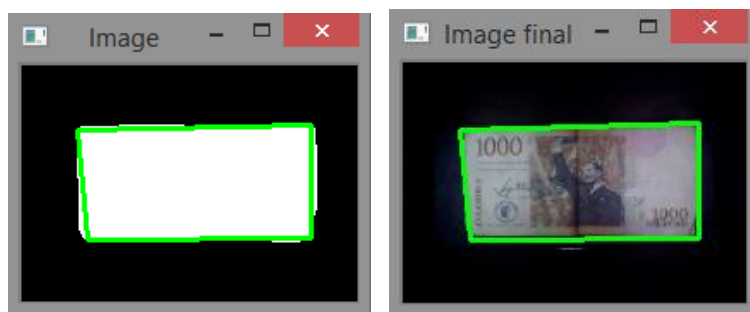


Ilustración 84. Figura rectangular encontrada en imagen binarizada e imagen original

Corrigiendo la perspectiva y se cambia el tamaño de la imagen, a un tamaño estandar para todas de 200 X 140 pixeles.



Ilustración 85. Imagen con perspectiva corregida

Esta imagen es una muestra de la salida del procesamiento, entrada para el inicio de la extracción de características.

Extracción de características

Para continuar con el proceso fue necesario revisar el estado del arte, libros y artículos, en estos textos, se encontró luego de analizarlos, que para hacer clasificación de objetos es posible extraer características de forma color, forma, textura, entre otros.

Al ser objetos con una forma similar se descartó este tipo de característica, teniendo como opciones las siguientes:

- **Promedio del valor de pixel en gris por filas**

Inicialmente la imagen se convierte a escala de grises



Ilustración 86. Imagen en gris

Se asigna un porcentaje de cuantas filas se quiere promediar, en este caso el valor es de 10% del valor de la altura de la imagen.

$$\text{Ancho de las filas} = t = 0.1 * h$$

$$\text{Ancho de las filas} = t = 0.1 * 140$$

Esto da un resultado 14 filas, lo que significa que se calculará el promedio de 14 filas

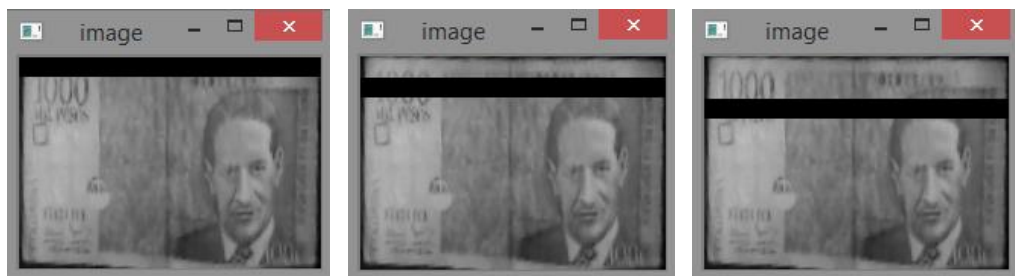


Ilustración 87. Recorrido de las filas en la imagen

En donde cada rectángulo de color negro de la imagen representa la posición de las 14 filas de la imagen que luego serán promediadas, y guardadas en un vector de características.

Como resultado el vector de características será de tamaño 10.

- **Promedio del valor de pixel en gris por bloques de filas y columnas**

Esta característica se calculó de forma similar a la anterior, sólo que se realiza un cálculo por bloques y no por filas

Inicialmente la imagen se convirtió a escala de grises



Ilustración 88. Imagen en gris

Se asigna un porcentaje de cuantas filas y columnas se quiere promediar, en este caso el valor es de 10% de la cantidad de valores a promediar de la altura y anchura de la imagen.

$$\text{Ancho de las filas} = tf = 0.1 * h$$

$$\text{Ancho de las columnas} = tc = 0.1 * w$$

$$\text{Ancho de las filas} = tf = 0.1 * 140 = 14$$

$$\text{Ancho de las columnas} = tc = 0.1 * 200 = 20$$

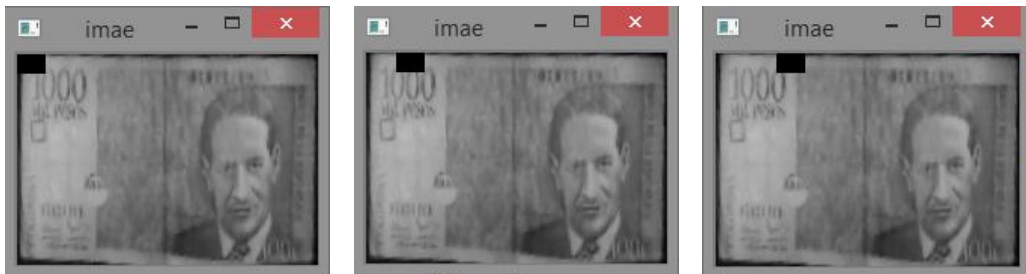


Ilustración 89. Recorrido de promedio por bloques

En donde cada rectángulo de color negro tendrá un tamaño de 14 X 20 pixeles, los cuales serán promediados y guardados en un vector de características.

Cada uno de los promedios se guardó en un vector de características arrojando un tamaño del vector de 100 características.

- **Promedio de histograma en RGB**

Una característica notable en algunos billetes, es su color, por la diversidad que presenta. Por esta razón se decidió trabajar con los valores de histograma de 3 canales.

Se dividió la imagen en valores de RGB como se puede ver en las siguientes figuras y se halló el histograma de cada uno de los canales

Los histograma presentan 256 valores, entonces se eligió un porcentaje que permitió el mismo tamaño para realizar el promedio el cual fue de 3.2 %

*Valores a promediar X histograma = Porcentaje de valores a promediar * ancho del vector*

$$\text{Valores a promediar} = 0.032 * 256 = 8$$

Dando como resultado un promedio de ocho valores para un histograma, los cuales fueron guardados en un vector de características.

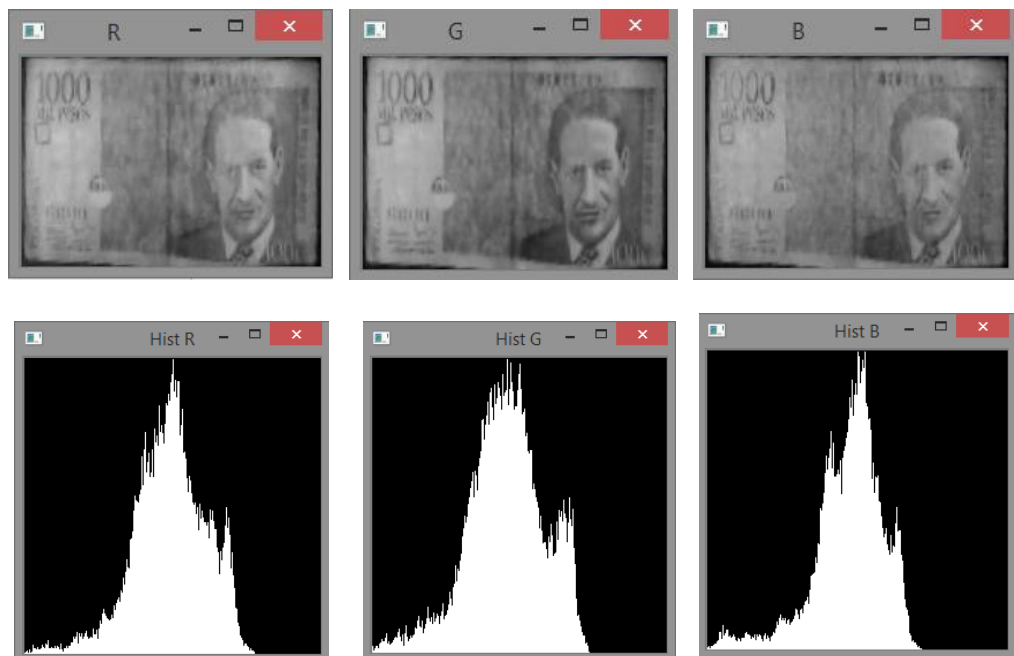


Ilustración 90. Imágenes en cada uno de los canales RGB / histograma de cada una de las imágenes

Cada canal presentaba un vector de características de 32 valores, al ser 3 histogramas el vector final de características compuesto por la unión de los 3 vectores arrojó como resultado un vector de 96 características.

- **Promedio de histograma en HUE**

El espacio de color HSV permite utilizar la característica del color directamente, sin necesidad de utilizar la cantidad de luz que hay en el entorno, esto se realiza calculando el

histograma del primer canal (HUE) para cada billete y guardando el promedio de cada 8 valores del mismo.

*Valores a promediar = Porcentaje de valores a promediar * ancho del vector*

$$\text{Valores a promediar} = 0.032 * 256 = 8$$

Dando como resultado un promedio de ocho valores para un histograma, los cuales fueron guardados en un vector de características.

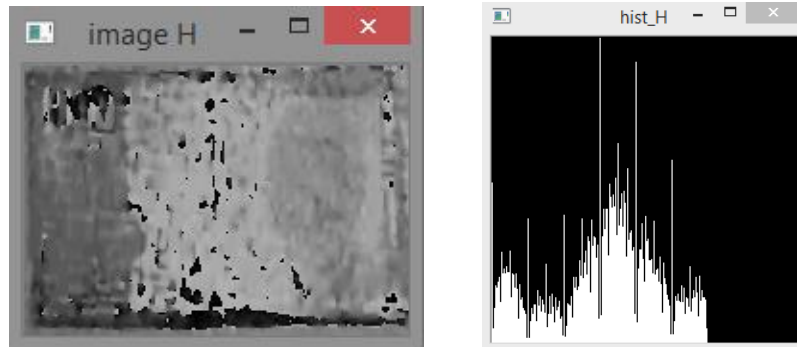


Ilustración 91. Imagen HUE / Histograma de imagen En HUE

El vector de características presenta un tamaño de 32 posiciones.

- **Local binary patterns (LBP)**

Esta es una característica de textura, invariante a la rotación, que analiza los vecinos alrededor dependiendo de cada pixel



Ilustración 92. Imagen en gris

El radio que se eligió para calcular los vecinos es 3:

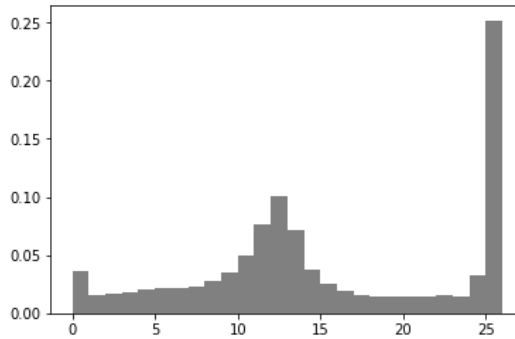


Ilustración 93. Histograma en LBP

El histograma arroja 26 posiciones, esto como resultado da un vector de 26 características, las cuales se guardan en un vector.

Almacenamiento de características

En Colombia se cuenta con 6 tipos de billetes para pronta recolección y 6 tipos de billete de nueva circulación.

Las clases se distinguen por letras, por ejemplo: cuando se habla de la clase “A”, se hace referencia al billete de \$.1000, en la tabla se pueden ver todas las clases con sus respectivos billetes, en donde la palabra “viejo” hace referencia a los billetes que esperan recoger y la palabra “nuevo” a los que están circulando en este momento.

Tabla 17. Clase con su respectivo billete

Clase	Billete	Estado
A	\$ 1.000	-
B	\$ 2.000	nuevo
C	\$ 2.000	viejo
D	\$ 5.000	nuevo
E	\$ 5.000	viejo
F	\$ 10.000	nuevo
G	\$ 10.000	viejo
H	\$ 20.000	nuevo
I	\$ 20.000	viejo

J	\$ 50.000	nuevo
K	\$ 50.000	viejo
M	No billete	-

Por facilidad de visualización y verificación de los datos se optó por utilizar un archivo de Excel para guardar los datos.

Se realizó un archivo donde se guardaron las características, lo que fue positivo porque minimizó tiempos en los cálculos con las imágenes, en la medida que se ejecutaba el programa. Este proceso permitió almacenar una sola vez las características de las imágenes, mientras que este archivo fue utilizado en todas las pruebas de entrenamiento, evitando operaciones innecesarias y redundantes. Es de resaltar que las características que se extraen, también son necesarias probarlas con el entrenamiento inicialmente, pero cuando ya se encuentra una buena respuesta, sólo se debe seguir actualizando el archivo a medida que aumentan las imágenes o las clases. Si se escogen unas muy buenas características que hagan distinguir los billetes entre ellos, no se necesitarán más el uso las imágenes para lo siguiente, sólo bastará con el archivo.

Entrenamiento

Asignarle un número a cada clase: la red neuronal necesita valores numéricos para realizar cálculos, es por eso que las clases (A,B,..., entre otras) que asignamos previamente a cada billete, debe cambiarse por un número que se llena en el vector de etiquetas en este caso es el vector "Y", por ejemplo: la clase A se le asigna el número 0, la clase B se le asigna el número 1 y así sucesivamente hasta terminar el número de clases, como se puede ver en la siguiente tabla.

Tabla 18. Clase con su respectiva etiqueta

Clase	Número
A	0
B	1
C	2
D	3
E	4
F	5
G	6
H	7
I	8
J	9

Se utilizó un Perceptron los siguientes

- Función de
- Número y capas de 100 de 50 y por

multicapa (MLP) como clasificador con hiperparámetros.

activación = 'tanh'
cantidad de capas ocultas = son 4 ocultas, la primera y segunda constan neuronas cada una, la segunda consta último la cuarta consta de 20 neuronas.

- La cantidad máxima de interacciones es 1000 y su tolerancia es de 0.0001, es decir, si la red neuronal no converge a un error menor o igual a 0.0001 en menos de 1000 iteraciones, entonces se guarda el último modelo.

K	10
M	11

Se realizaron varias pruebas con las características que previamente se explicaron

Como prueba inicial se hizo una captura de 100 imágenes por clase, con 3 tipos de características, arrojando resultados de exactitud no mayor al 85% en el entrenamiento, como se puede ver a continuación.

Tabla 19. Porcentaje de exactitud por tipo de característica

Tipo de características	Porcentaje de exactitud
Promedio del valor de pixel en gris por filas	70%
Promedio del valor de pixel en gris por bloques de filas y columnas	70%
Histograma RGB	85%

Al obtener resultados poco exitosos, se decidió aumentar el número de imágenes; esta vez realizando la prueba solamente con el tipo de características del histograma en RGB, ya que fueron los que alcanzaron mejores niveles. Al aumentar las imágenes de 100 a 400 se consiguió aumentar el porcentaje de exactitud de un 85% a un 92%, resultados satisfactorios para una primera aproximación.

De este modo se decidió implementar y realizar pruebas, pero cuando se intentaba reconocer billetes no entrenados, el avance no era significativo; por este motivo se decidió aumentar la cantidad de imágenes a 700 por billete, logrando así una exactitud de 95%. Al implementar inmediatamente este nuevo modelo, mejoró considerablemente el reconocimiento de los billetes.

Otro obstáculo fue el uso continuo por el exceso en la realización de pruebas, que ocasionó que el sistema comenzara a fallar; pues la intensidad de luz que proporcionaban los LED's, disminuían el voltaje de la batería.

En este orden de ideas, con base en pruebas anteriores, se observó que aumentar el número de imágenes mejoraba los resultados; por ello se optó por realizar un aumento de datos sin tener que capturar más imágenes de nuevo, efectuando una serie de modificaciones a algunas imágenes como rotarlas, voltearlas de manera horizontal y vertical, desenfocarlas, aplicarles ruido, ya sea gaussiano o de impulso (sal y pimienta); generando un aumento de imágenes del 40% aproximadamente, con respecto al que se tenía anteriormente. Se almacenaron alrededor de 1200 imágenes por cada clase; la red

proporcionó una exactitud del 96%, pero aunque el modelo en RGB mejoró la falla que presentaba, varios de los resultados no eran los esperados, obligando a extraer características más robustas.

La escala de color en RBG tiende a fallar por la iluminación debido que esta hace cambiar drásticamente los valores del histograma; entonces se eligió una escala de color diferente, que estuviera mejor relacionada con el color evitando utilizar la cantidad de luz que hay en ella, de este modo, se extrajo características de histograma de sólo el valor del primer canal (HUE), ya que los otros canales determinan la luminosidad (Saturation) y la cantidad de negro que hay en la imagen (Value). Para estas características se logró un resultado del 95 % de exactitud y al momento de implementarse presentaba niveles óptimos con billetes diferentes a los proporcionados en el entrenamiento.

Adicional a esto, se decidió aumentar la robustez del sistema adicionándole otro tipo de característica diferente al color, teniendo en cuenta que hay billetes con colores muy similares que tienden a confundirse entre ellos, aplicando la característica de textura LBP, con estos dos tipos de características se obtuvo un 96% de exactitud.

El proyecto avanzó significativamente cuando se cambiaron los tipos de características y se aumentaron los datos. Los tipos de características finales con mejores resultados son:

1. Valores del histograma en el primer canal (HUE) de la escala HSV.
2. Valores histograma con local binary patterns.

Finalmente, los dos tipos de características arrojaron un vector de 58 características que ingresan al entrenamiento.

3.1.2 Construcción prototipo 2

Partiendo del conocimiento adquirido del primero prototipo, se planteó construir otro, con las mismas consideraciones que el prototipo anterior consolidando una base de datos con las imágenes en diferentes ambientes, para luego realizar diferentes pre procesados que dependen del diseño físico del guante.

- **Diseño físico 1**

El objetivo inicial de pre procesamiento es reconocer o localizar el guante para luego segmentar el billete contenido en éste.

Inicialmente se toma la imagen a color (escala RGB), Para luego pasar a una escala de HSV, para trabajar específicamente con los colores

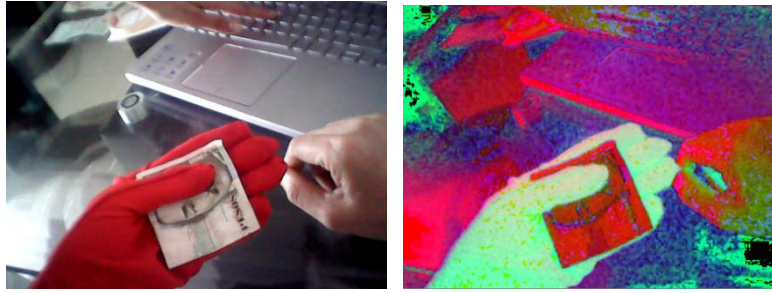


Ilustración 94. Imagen color / imagen HSV

Luego se binariza con los rangos pertenecientes al color rojo y se le aplica una dilatación



Ilustración 95. Imagen binarizada

Se recorta la imagen del tamaño del contorno que contenga mayor área, después se recorre de manera horizontal de derecha a izquierda y viceversa; de arriba a abajo y viceversa, la imagen siguiente muestra cómo de derecha a izquierda recorre el guante dibujando el contorno.

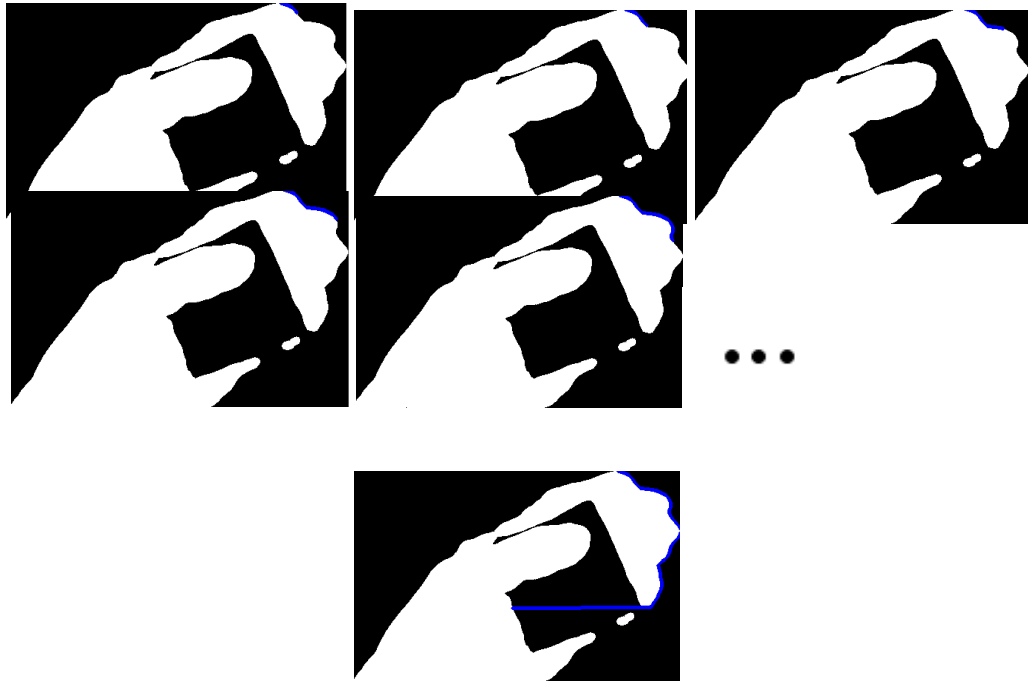


Ilustración 96. Recorrido del guante

El contorno se forma buscando el cambio de negro a blanco, el primer cambio que se encuentre se determina como primer punto, luego se pasa a la siguiente fila, y se sigue recorriendo hasta encontrar otro cambio (punto 2). Si la distancia entre esos dos puntos es mayor a un valor establecido, significa que existe la posibilidad de que la imagen del guante no esté completa, buscando entre los puntos de mayores distancias, en cada uno de los diferentes recorridos, cerrar los huecos que posee el guante, como se puede ver a continuación:

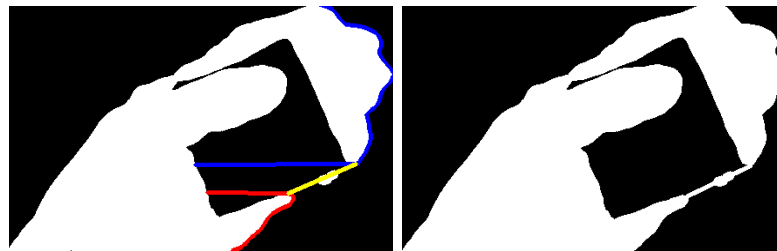


Ilustración 97. Imagen a color con los huecos cerrados / imagen a blanco y negro con los huecos cerrados

Luego de cerrar los huecos y completar el contorno se invierten los colores de la imagen, y se pinta de negro lo que esté por fuera del guante, como se puede ver a continuación:

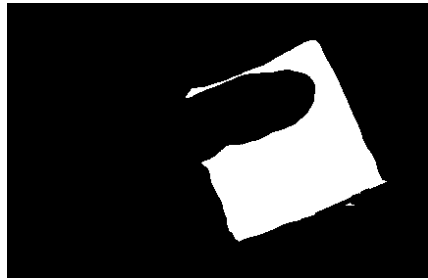


Ilustración 98. Imagen invertida y rellena externamente

Finalmente se recortan la imagen original con los puntos del contorno previamente hallado y se une con la imagen anterior, en donde se puede visualizar únicamente el billete, como se puede ver a continuación:



Ilustración 99. Imagen recortada y combinada con la original

Para el primer diseño se realizaron varias capturas de imágenes aproximadamente 400 de cada billete, en diferentes ambientes, con diferente iluminación. Este guante presentó varios inconvenientes en algunas imágenes a la hora de pre procesar, como por ejemplo.





Ilustración 100: Problemas en el pre procesado de las imágenes del diseño prototipo 1

Como se puede apreciar, las imágenes anteriores presentan muchos problemas a la hora de segmentar el billete, puesto que al realizar el pre procesado algunas de ellas presentan dificultades para localizar la parte interna del guante, que en este caso sería el billete.

Para las imágenes que tuvieron un adecuado pre procesamiento se realizó una extracción de características similar del prototipo 1 y se entrenó con estas características obteniendo un porcentaje de exactitud en el cross validation no mayor al 85%

La cantidad de imágenes que se lograban pre procesar adecuadamente eran aproximadamente el 60% del total de las muestras, cantidad insuficiente para realizar el entrenamiento y obtener un buen modelo y por ende, realizar la predicción. Además, muchas de las imágenes estaban mal pre procesadas; lo que significaba, que de las pocas imágenes que quedaban, algunas características no eran las correctas.

Otro de los problemas que se presentó, fue la manera en la que la persona tomaba el billete, esto generaba distorsión en la imagen ya que el pulgar hacía que se perdiera información del objeto, adicionalmente, algunas de las imágenes no estaban bien segmentadas, pudiéndose ver el fondo o sólo una porción del billete, como se puede ver a continuación



Ilustración 101. Imágenes mal segmentadas

Al final no fue posible realizar pruebas de predicción con este diseño puesto que no presentaba buenos resultados. Debido a esto se continuó con el diseño 2.

- **Diseño físico 2**

En este diseño presenta leves modificaciones del diseño anterior, puesto que se notó que el pulgar hacia que se perdiera información del billete y la segmentación en varias ocasiones presentaba malos resultados, para ello se optó por adicionar un bolsillo al borde de la palma de la mano, para guardar el billete al interior del guante.

Para extraer el billete del guante se realizan dos pre procesados, el primero es para extraer el guante del fondo y el siguiente es para extraer el billete del guante.

El primer pre procesado se realizó de la siguiente manera: a partir de imagen a color (RGB), se realizó un posterizado:



Ilustración 102. Imagen original / imagen posterizada

Esa imagen posterizada se pasa a escala de HSV como lo muestra la siguiente imagen:



Ilustración 103: imagen posterizada en escala HSV

Luego se procede binarizar el color rojo, para obtener una imagen como lo muestra la siguiente figura, en donde se puede ver el guante de color blanco. En este paso se analizan los contornos al interior del guante; si estos valores cumplen con un área previamente definida, significa que se ha encontrado el billete.



Ilustración 104. Imagen binarizada

Se recorta la imagen con un contorno que sea de mayor área a la del billete original, para generarle un fondo, como se puede ver a continuación:



Ilustración 105. Imagen original recortada con los puntos del contorno interior del guante

Luego se hace otro pre procesado (Ver pre procesado Prototipo 1) para segmentar el billete de la siguiente manera:



Ilustración 106. Imagen Pre procesada nuevamente con el código del prototipo 1

El primer diseño se modificó añadiéndole un bolsillo en la parte inferior de la palma de la mano, para delimitar que el billete siempre esté en una misma posición, con el dobléz en la parte externa, además se unieron el pulgar y el índice acemejado una pata palmeada, como se observa en la siguiente figura:

Para este diseño se comenzaron capturando alrededor de 300 imágenes de cada billete en diferentes ambientes.

A la hora de realizar el pre procesamiento, el segundo diseño presentó mejores resultados con respecto al primero, segmentando el billete de una mejor manera, gracias a que se utilizaron varios métodos iniciando con la segmentación inicial del guante y a partir de este realizar otra segmentación para obtener el billete.

En algunas ocasiones para segmentar el color rojo del guante cuando había mucha iluminación presentaba colores que tendían a ser blancos y cuando había poca iluminación presentaba colores negros, generando aberturas en el guante como se puede ver a continuación



Ilustración 107: Imágenes de guantes incompletos

Para ello se decidió ampliar el rango de rojo para que obtuviera pudiera binarizar el color rojo en todos sus tonos. Al realizar esto se notó, que en algunos billetes predominaba el color rojo, como es el caso del billete de \$10.000 de vieja generación, impidiendo la adecuada segmentación de este.



Ilustración 108: Billeto de \$10.000 binarizado

Se decidió realizar la extracción de características a partir de las imágenes segmentadas adecuadamente, extrayendo las mismas características que se extrajeron para el prototipo 1, las cuales fueron LBP y promedio histogramas en HUE.

Luego de extraer las características y guardarlas en un archivo, se procedió a realizar el entrenamiento, el cual presentó una exactitud en el cross validation del 92%, sin embargo, se pudo ver que la matriz de confusión que presentaba tenía valores muy desiguales, esto debido a que algunas imágenes no segmentaron bien, entonces no había la misma cantidad de imágenes en todas las clases, lo que significa que se aprendió mejor unas clases que otras.

Se realizaron varias pruebas en diferentes entornos con el modelo que presentó el entrenamiento, la predicción inicialmente se realizó con las imágenes que se

ingresaron al modelo, presentando excelentes resultados, sin embargo al implementarse en tiempo real los resultados no fueron los esperados, además de lo mencionado anteriormente, también influían variables que no podían ser controladas, como son la iluminación y la distancia a la cámara, las cuales sí pudieron controlarse en el prototipo 1, debido a esto se puede asegurar que el modelo estaba sobre ajustado, ya que presentaba muy buenos resultados con los datos que se entrenó, pero presentaba resultados desfavorables con otras imágenes.

Algunas pautas que pueden mejorar la predicción son:

- Aumento de imagen, ya sea capturando más imágenes y/o realizando un aumento de datos como se explicó previamente
- Asegurando una cantidad de imágenes iguales a la hora de realizar el pre procesado.
- Aplicar otras técnicas que permitan el uso de imágenes sin tanto pre procesamiento

Tomando en cuenta las consideraciones previas se realizó un tercer prototipo.

• **Diseño físico 3**

Para el diseño físico 3 se planteó utilizar un guante de un color que no predominara en los billetes como es el negro, que tuviera un contorno verde alrededor de este, el cual evitara que objetos externos al guante entraran en los rangos de color negro separando el guante del fondo.

El pre procesado que se realizó en esta sección, es con el objetivo de recortar la parte de la imagen donde se encuentra el guante, para reducir el ruido exterior y mejorar el modelo. El pre procesado se realizó de la siguiente manera

A partir de la imagen a color (RGB)



Ilustración 109. Captura de imagen diseño físico 3

Se realizan dos partes

- Se pasa la imagen a color a escala de grises para binarizar la imagen aplicando técnicas de umbralización.



Ilustración 110. Imagen en gris / imagen binarizada

Por otro lado, se posteriza la imagen original para luego pasarla a escala de HSV y extraer el contorno verde del guante, binarizando este color con sus rangos establecidos.

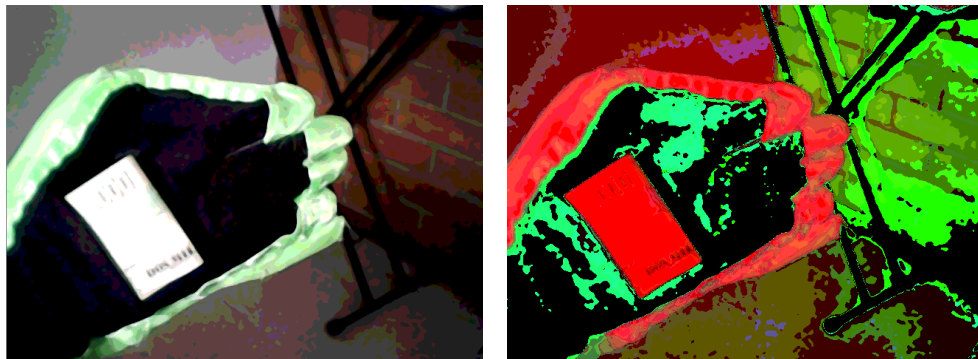


Ilustración 111. Imagen binarizada con rangos del color verde e invertida

- Se unen las dos imágenes binarizadas por medio de un and lógico



Ilustración 112. Unión entre las dos imágenes binarizadas

Se encuentra el contorno de mayor área que presenta un contorno interno con un área interna predefinida

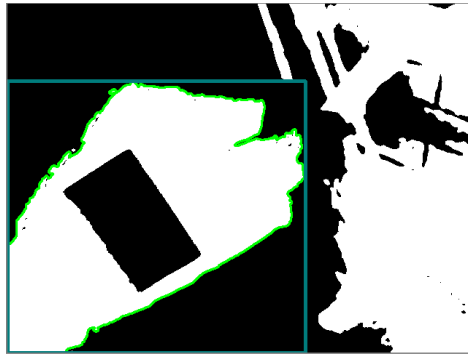


Ilustración 113. Contorno guante enmarcado en un rectángulo

Para finalmente ser recortado:



Ilustración 114. Imagen original recortada

Se reduce a un tamaño estándar, para que todas las imágenes tengan el mismo tamaño



Ilustración 115. Cambio de tamaño de imagen recortada

Esta imagen se lleva a las siguientes etapas, en las cuales la red neuronal realiza internamente los procesos (pre procesar y extraer características) necesarias para reconocer la denominación del billete.

Se comenzaron haciendo capturas de 500 imágenes por billete y realizado el pre procesado del guante como se mencionó anteriormente.

Al realizar el pre procesado varias imágenes presentaban problemas como que el guante no se cerraba completamente, representando un 10% de imágenes segmentadas del billete que se perdían.

A partir de esto, se extrajeron características y realizó el entrenamiento de la red, obteniendo un resultado de exactitud del 95%, a partir de este modelo se realizó la predicción, que mostró resultados similares al anterior, en donde se acertó correctamente las imágenes de los billetes con los que se entrenó, pero fracasó al no reconocer imágenes nuevas que se tomaban en tiempo real.

Siguiendo con las pruebas, se decidió aumentar la cantidad de imágenes a 1000 por billete, con el objetivo de mejorar el modelo, volviendo así a realizar el entrenamiento. La exactitud del cross validation aumentó a 97%. Al realizar pruebas con este nuevo modelo se obtuvo mejores resultados que al tener menos imágenes, permitiendo así la predicción en algunos billetes, sin embargo aunque realizaba algunas predicciones bien, otras las realizaba erróneamente.

Por este motivo se decidió realizar el diseño del prototipo 3 realizando técnicas diferentes, como es Deep learning.

- Deep learning

A partir de las imágenes con el guante pre procesado se decidió realizar otras técnicas como Deep learning para el reconocimiento de los billetes, para ello además de utilizar las imágenes anteriores, luego se separan las imágenes en imágenes de para el entrenamiento e imágenes para el test, y antes de realizar el entrenamiento se realizó un aumento de datos, aumentando el brillo de la imagen, agregando ruido, desenfoquándola, rotándola, entre otras, para darle mayor robustez a la red. Finalmente donde el modelo se diseñó con los siguientes hiperparámetros:

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 96, 96, 32)	896
activation_1 (Activation)	(None, 96, 96, 32)	0
batch_normalization_1 (Batch Normalization)	(None, 96, 96, 32)	384
max_pooling2d_1 (MaxPooling2D)	(None, 32, 32, 32)	0
dropout_1 (Dropout)	(None, 32, 32, 32)	0
conv2d_2 (Conv2D)	(None, 32, 32, 64)	18496
activation_2 (Activation)	(None, 32, 32, 64)	0
batch_normalization_2 (Batch Normalization)	(None, 32, 32, 64)	128
max_pooling2d_2 (MaxPooling2D)	(None, 16, 16, 64)	0
dropout_2 (Dropout)	(None, 16, 16, 64)	0
conv2d_3 (Conv2D)	(None, 16, 16, 64)	36928
activation_3 (Activation)	(None, 16, 16, 64)	0
batch_normalization_3 (Batch Normalization)	(None, 16, 16, 64)	64
max_pooling2d_3 (MaxPooling2D)	(None, 8, 8, 64)	0
dropout_3 (Dropout)	(None, 8, 8, 64)	0
conv2d_4 (Conv2D)	(None, 8, 8, 128)	73856
activation_4 (Activation)	(None, 8, 8, 128)	0
batch_normalization_4 (Batch Normalization)	(None, 8, 8, 128)	32
max_pooling2d_4 (MaxPooling2D)	(None, 4, 4, 128)	0
dropout_4 (Dropout)	(None, 4, 4, 128)	0
flatten_1 (Flatten)	(None, 2048)	0
dense_1 (Dense)	(None, 200)	409800
activation_5 (Activation)	(None, 200)	0
batch_normalization_5 (Batch Normalization)	(None, 200)	800
dropout_5 (Dropout)	(None, 200)	0
dense_2 (Dense)	(None, 11)	2211
=====		
Total params: 543,595		
Trainable params: 542,891		
Non-trainable params: 704		

Ilustración 116. Resumen del diseño de la red neuronal

La API de Keras, permite una visualización de la secuencia del diseño de la red, en donde se describe las salidas de cada función que se utilizó, el Anexo 7 presenta un detalle de las entradas y salidas más detalladamente.

Para compilar el modelo se utilizó el optimizador Adam en donde se definen variables como:

- la tasa de aprendizaje= 0.003
- Pérdida = “categorical_crossentropy”

Finalmente se entrenó con los parámetros predichos arrojando resultado de entrenamiento del 95%, que a la hora de realizar la pruebas con la predicciones, arroja un buen resultado de exactitud en comparación con los anteriores prototipos; sin embargo, existe un margen de error, esto se debe a la iluminación, distancia de la cámara hacia el guante, entre otras, estos errores pretenden minimizarse a futuro generando una red más robusta con otros métodos, en donde aprenda mejor las características. La red también puede mejorar aumentando el número de imágenes, ya que las imágenes entran a la red sin pre procesar.

3.2 RENDIMIENTO DE LA RED

Para verificar cómo cambiaba el rendimiento de la red a medida que aumentaba las clases, se comenzó entrenando con dos de ellas para ir aumentando gradualmente de a dos características, hasta llegar a la cantidad de billetes que existen.

Tabla 20. Rendimiento de la red

Número de clases	Rendimiento Red
2	99.80%
4	98.00%
6	97.40%
8	97.10%
10	97.00%
12	96.60%

En la gráfica se puede ver la disminución de la validación cruzada a medida que se aumenta el número de clases:

Diminución de rendimiento por número de clases

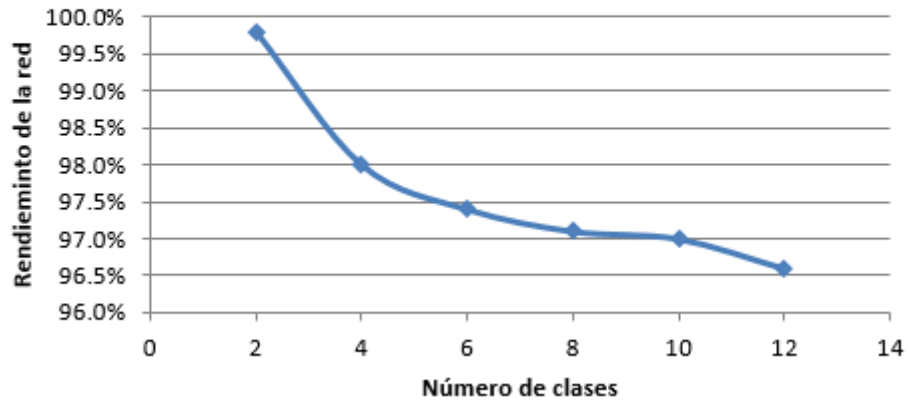


Ilustración 117. Disminución rendimiento por número de clases

Con respecto al funcionamiento del sistema controlado, para ser el primer prototipo, se comprobó que el algoritmo reconoce satisfactoriamente los billetes, sin embargo, existe un margen de error mínimo en el reconocimiento de estos, para ello se recomienda realizar varias predicciones con diferentes posiciones del billete para mejor clasificación.

Como futuros desarrollos se pretende mejorar la precisión del sistema generando robustez en los algoritmos.

4. CONCLUSIONES

Definir los billetes de mayor circulación y los patrones característicos de cada uno de ellos.

- Cabe resaltar que los diferentes textos consultados fueron de gran ayuda durante el proyecto, en estos se identificaron las diferentes técnicas que permitieron encontrar patrones característicos que diferenciaban los 11 tipos de billetes procesados por el sistema.
- Dentro de este mismo objetivo se resalta la importancia de contar con un buen sistema de iluminación en entornos controlados debido a que es uno de los factores que más inciden en el procesamiento, y por tanto necesita de una iluminación óptima, para garantizar la obtención de los resultados deseados a la hora de extraer los patrones característicos.

Para el objetivo de Implementar algoritmos de visión artificial para reconocimiento y clasificación de objetos y verificar el funcionamiento del sistema.

- Para a obtener una mayor confiabilidad del sistema, y evitar el sobre ajuste, es necesario aumentar la cantidad de imágenes en diversos ambientes y escoger las mejores características para la aplicación.
- Se evidencio que las características de la cámara y el tipo de iluminación que se utilizó influye a la hora de realizar pruebas, para el experimento 1 y 2 es necesario utilizar la misma cámara con la que se capturaron las imágenes que se utilizaron en el entrenamiento y para el experimento 3, puesto que Deep-learning extrae sus propias características es posible variar la cámara, siempre y cuando haya diversidad de imágenes con diferentes cámaras.
- Como mejoras para el experimento 3, es posible utilizar otras técnicas avanzadas de Deep-learning como lo es el transfer-learning, el cual se utiliza cuando no hay una cantidad de datos significativos para el proceso, puesto que permite a la red aprender a extraer características; esto lo realiza por medio de modelos previamente entrenados con millones de imágenes, que luego se utilizan eliminando la última capa y volviendo a entrenar la red con las imágenes personalizadas.
- Se pudo evidenciar que realizar entrenamiento con una cantidad exagerada de hiperparámetros (iteraciones, épocas, tasa de aprendizaje, número de neuronas y capas ocultas, entre otros) no suele dar mejor efectividad y genera el sobre ajuste de la red, se recomienda comenzar con modelos pequeños y variar los hiperparámetros, a medida que se vea la efectividad de ellos, aumentarlos gradualmente, hasta obtener un modelo robusto.
- Es necesario crear algoritmos que generen una pequeña distorsión en las imágenes con el objetivo de aumentarlas, debido a dificultad de adquirir una cantidad de imágenes significativa para el entrenamiento de la red.

- Para utilizar técnicas de deep-learning, se recomienda usar equipos con altas características puesto que se deben realizar gran cantidad de cálculos, en internet existen varios recursos que permiten el uso de esas maquina súper potentes sin necesidad de comprar una, una opción gratuita es utilizar Google Colaboratory, el cual es un servicio de Google que permite desarrollar algoritmos en Python con el uso de GPU con altas prestaciones. (Kdnugget, s.f)

Integrar los algoritmos en un sistema embebido.

- Se logra evidenciar que la integración de los algoritmos en los sistemas embebidos genera un cambio en la velocidad de procesamiento, para el algoritmo ejecutado en una computadora, se demora un tiempo aproximado de 1.5s de ejecución del código, mientras en la Raspberry se demora 2.2s
- Se debe tener en cuenta que la versión de librerías que se utilicen deben ser igual a la que se instaló para realizar el entrenamiento.
- Como trabajos futuros se pretende
 - Para el experimento 1, encontrar la cámara y la iluminación adecuada que permita resaltar mejor las características de los billetes.
 - Para el experimento 3 utilizar técnicas más avanzadas como transferlearning que permitan mejorar la efectividad en predicción del mismo.
 - Desarrollar la aplicación en un dispositivo móvil android y/o ios debido a que son dispositivos de uso frecuente en las personas con discapacidad visual.

5. REFERENCIAS

- Academic. (s.f.). *Esacademic.com*. Recuperado el 14 de Mayo de 2018, de Visión artificial: <http://www.esacademic.com/dic.nsf/eswiki/287535>
- Almudena. (s.f.). *Twenergy*. Recuperado el 17 de Mayo de 2018, de LOS INVENTORES DE LA LUZ LED AZUL GANAN EL PREMIO NOBEL DE FÍSICA 2014: <https://twenergy.com/a/los-inventores-de-la-luz-led-azul-ganan-el-premio-nobel-de-fisica-2014-1427>
- Amazon. (16 de Octubre de 2016). <https://www.amazon.com>. Obtenido de <https://amzn.to/2RmRj1H>
- Amazon. (28 de Abril de 2016). <https://www.amazon.es>. Obtenido de <https://amzn.to/2zbffxH>
- Arquys. (Diciembre de 2012). Luz de halógeno. *Revista ARQHYS*. Recuperado el 18 de Diciembre de 2017, de <http://www.arqhs.com/construccion/luz-halogeno.html>.
- Banco de la República - Colombia. (s.f.). *Billetes y monedas*. Recuperado el 5 de Mayo de 2018, de [Banrep.gov.co](http://www.banrep.gov.co): <http://www.banrep.gov.co/es/billetes-y-monedas>
- Banco de la República- Colombia. (2013). *Marca táctil en los billetes*. Recuperado el 5 de Mayo de 2018, de [Banrep.gov.co](http://www.banrep.gov.co): <http://www.banrep.gov.co/es/node/31529>
- Banco de la República Colombia. (s.f.). *Preguntas y respuestas - Nueva familia de billetes*. Recuperado el 5 de Mayo de 2018, de [Banrep.gov.co](http://www.banrep.gov.co): <http://www.banrep.gov.co/es/nueva-familia-billetes-preguntas>
- BBVA open mind. (15 de Abril de 2015). *OpenMind*. Recuperado el 18 de Diciembre de 2017, de 5 ejemplos de tecnología para personas ciegas: más allá del Braille: <https://www.bbvaopenmind.com/tecnologia-para-invidentes-mas-alla-del-braille/>
- Bcnvision. (11 de Abril de 2017). *Bcnvision.es*. Recuperado el 16 de Mayo de 2018, de Sistemas de iluminación para aplicaciones de visión artificial (Parte II): <http://www.bcnvision.es/blog-vision-artificial/iluminacion-vision-artificial2/>
- Blanco Romero, E. (2015). *Plan de investigación*. Barcelona: Universidad politécnica de cataluña. Recuperado el 22 de Mayo de 2018, de <file:///D:/Downloads/PR%20Blanco%20Romero,%20Elena.pdf>
- Boudeguer Simonetti, A., Prett Webe, P., & Squella Fernández, P. (Octubre de 2010). *Manual de Accesibilidad Universal*. Recuperado el 05 de Octubre de 2018, de https://www.mutual.cl/Portals/0/PDF/dividendo_social/manual_accesibilidad_OK_sello_baja.pdf
- Brownlee, J. (3 de Julio de 2017). *Gentle Introduction to the Adam Optimization Algorithm for Deep Learning*. Recuperado el 22 de Ocutubre de 2018, de

machinelearningmastery.com: <https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/>

Budhiraja, A. (15 de Diciembre de 2016). *Dropout in (Deep) Machine learning*. Recuperado el 22 de Octubre de 2018, de medium.com: <https://medium.com/@amarbudhiraja/https-medium-com-amarbudhiraja-learning-less-to-learn-better-dropout-in-deep-machine-learning-74334da4bfc5>

CEADE. (6 de Mayo de 2015). *Luminotecniaceade.blogspot.com.co*. Recuperado el 17 de Mayo de 2018, de Iluminación por fibra óptica: <http://luminotecniaceade.blogspot.com.co/2015/05/iluminacion-por-fibra-optica.html>

Clínica RAHHAL. (s.f.). *rahhal.com*. Recuperado el 17 de Mayo de 2018, de Mácula ocular: ¿qué es y qué función desempeña?: <https://www.rahhal.com/blog/macula-ocular-que-es-y-que-funcion-desempena/>

Definicion ABC. (s.f.). *Definicionabc.com*. Recuperado el 14 de Mayo de 2018, de Definición de Imagen Digital: <https://www.definicionabc.com/tecnologia/imagen-digital.php>

Definicion ABC. (s.f.). *Definicionabc.com*. Recuperado el 14 de Mayo de 2018, de Definición de Cámara Fotográfica: <https://www.definicionabc.com/tecnologia/camara-fotografica.php>

Del Valle Hernández, L. (s.f.). *Detector de bordes Canny cómo contar objetos con OpenCV y Python*. Recuperado el 24 de Octubre de 2018, de programarfácil.com: <https://programarfácil.com/blog/vision-artificial/detector-de-bordes-canny-opencv/>

DocShop. (15 de Septiembre de 2014). *Docshop.com*. Recuperado el 18 de Mayo de 2018, de Acerca del ojo: <https://www.docshop.com/es/education/vision>

Doukkali, F. (20 de Octubre de 2017). *Batch normalization in Neural Networks*. Recuperado el 22 de Octubre de 2018, de medium.com: <https://towardsdatascience.com/batch-normalization-in-neural-networks-1ac91516821c>

EcuRed. (s.f.). *Vision Artificial*. Recuperado el 20 de Mayo de 2018, de Ecured.cu: http://www.ecured.cu/index.php/Visión_Artificial?PageSpeed=noscript

El país. (01 de Noviembre de 2013). Colombia tiene 1,14 millones de personas con problemas visuales. *El país*. Recuperado el 11 de Noviembre de 2017, de <http://www.elpais.com.co/colombia/tiene-1-14-millones-de-personas-con-problemas-visuales.html>

El Tiempo. (14 de Julio de 2015). La falta de visión no fue un impedimento. *El Tiempo*. Recuperado el 15 de Diciembre de 2017, de <http://www.eltiempo.com/archivo/documento/CMS-16094395>

El Tiempo. (22 de Mayo de 2017). La deuda del país con su población en condición de discapacidad. Recuperado el 6 de Mayo de 2018, de

<http://www.eltiempo.com/colombia/otras-ciudades/poblacion-en-condicion-de-discapacidad-y-sus-necesidades-en-colombia-90880>

Elío, J. (18 de Enero de 2016). *El accesorio más loco para tu teléfono: Un endoscopio para Android*. Recuperado el 23 de Octubre de 2018, de elandroidelibre.espanol.com: <https://elandroidelibre.espanol.com/2016/01/endoscopio-para-android.html>

Family connect. (s.f.). *Discapacidad visual: Visión general*. Obtenido de [familyconnect.org](http://www.familyconnect.org): <http://www.familyconnect.org/info/despues-del-diagnostico/trabajar-con-los-profesionales-medicos/discapacidad-visual/135>

FotoNostra. (s.f.). *Fotonostra.com*. Recuperado el 17 de Mayo de 2018, de Las Lentes del Objetivo: <http://www.fotonostra.com/fotografia/lentesobjetivo.htm>

francobordo. (s.f.). *Luz Fluorescente LED Hermetica Version en Angulo*. Recuperado el 17 de Mayo de 2018, de <https://www.francobordo.com/luz-fluorescente-led-hermetica-version-angulo-p-334396.html>

García Moreno, C. (11 de 22 de 2016). *¿Qué es el Deep Learning y para qué sirve?* Recuperado el 22 de Octubre de 2018, de [Indracompany.com](http://www.indracompany.com): <https://www.indracompany.com/es/blogneo/deep-learning-sirve>

Garcia, I. B. (11 de Mayo de 2018). *Camaras.uno*. Recuperado el 17 de Mayo de 2018, de Cómo elegir y comprar una buena cámara fotográfica digital en 2018: <https://camaras.uno/guia-para-comprar-una-camara-fotografica-caracteristicas-tipos-de-camara/>

GE Intelligent Platforms. (s.f.). *Classification of objects from video streams*. Obtenido de https://www.wb-ip.com.au/uploads/3/8/8/4/38841341/classification-of-objects-from-video-streams-wp-gft887_0.pdf

Gestion Global de Recursos S.L. (s.f.). *globalexchange.es*. Recuperado el 14 de Mayo de 2018, de Orígenes e historia del peso colombiano: <https://www.globalexchange.es/monedas-del-mundo/peso-colombiano>

Gomez, F. (23 de Marzo de 2015). *deustoformacion.com*. Recuperado el 19 de Mayo de 2018, de 8 características de la imagen digital que debes conocer: <https://www.deustoformacion.com/blog/disenio-produccion-audiovisual/8-caracteristicas-imagen-digital-que-debes-conocer>

Google . (s.f.). *ML Practicum: Image Classification*. Obtenido de developers.google.com: <https://developers.google.com/machine-learning/practica/image-classification/convolutional-neural-networks?hl=es-419>

Google Cloud. (s.f.). *Intro to TensorFlow*. Obtenido de [coursera.org](https://www.coursera.org): <https://www.coursera.org/learn/intro-tensorflow>

Google. (s.f.). *Estudio detallado del AA: Entrenamiento y pérdida*. Recuperado el 22 de Octubre de 2018, de developers.google.com:

<https://developers.google.com/machine-learning/crash-course/descending-into-ml/training-and-loss?hl=es-419>

Google. (s.f.). *Glosario sobre aprendizaje automático*. Recuperado el 22 de Octubre de 2018, de https://developers.google.com/machine-learning/crash-course/glossary?hl=es-419#gradient_descent

Google. (s.f.). *Reducción de la pérdida: Tasa de aprendizaje*. Recuperado el 22 de Octubre de 2018, de [developers.google.com: https://developers.google.com/machine-learning/crash-course/reducing-loss/learning-rate?hl=es-419](https://developers.google.com/machine-learning/crash-course/reducing-loss/learning-rate?hl=es-419)

Gracia, L. M. (9 de Octubre de 2013). *¿Qué es OpenCV?* Recuperado el 20 de Mayo de 2018, de [Unpocodejava.com: https://unpocodejava.com/2013/10/09/que-es-opencv/](https://unpocodejava.com/2013/10/09/que-es-opencv/)

Hisystem. (s.f.). *hisystem.com.ar*. Recuperado el 19 de Mayo de 2018, de http://www.hisystem.com.ar/2011/tecnologia_camcctv.html

Imnovation. (s.f.). *Aprendizaje reforzado: cuando las máquinas*. Recuperado el 16 de Octubre de 2018, de [Imnovation-hub.com: https://www.imnovation-hub.com/es/transformacion-digital/aprendizaje-reforzado-cuando-las-maquinas-aprenden-solas/](https://www.imnovation-hub.com/es/transformacion-digital/aprendizaje-reforzado-cuando-las-maquinas-aprenden-solas/)

Infaimon. (s.f.). *Infaimon.com*. Recuperado el 17 de Mayo de 2018, de Fibra Óptica: <https://www.infaimon.com/es/fibra-optica>

Inteldig. (7 de Mayo de 2018). *TUTORIAL TENSORFLOW: CÓMO COMENZAR CON TENSORFLOW, REQUISITOS Y EJEMPLOS*. Obtenido de [Inteldig.com: https://www.inteldig.com/2018/05/tutorial-tensorflow-como-comenzar-con-tensorflow-requisitos-y-ejemplos/](https://www.inteldig.com/2018/05/tutorial-tensorflow-como-comenzar-con-tensorflow-requisitos-y-ejemplos/)

Kdnugget. (s.f.). *Deep Learning Development with Google Colab, TensorFlow, Keras & PyTorch*. Obtenido de [kdnugget.com: https://www.kdnuggets.com/2018/02/google-colab-free-gpu-tutorial-tensorflow-keras-pytorch.html](https://www.kdnuggets.com/2018/02/google-colab-free-gpu-tutorial-tensorflow-keras-pytorch.html)

Keras. (s.f.). *Keras: The Python Deep Learning library*. Recuperado el 22 de Octubre de 2018, de [keras.io: https://keras.io/](https://keras.io/)

Kréalo. (04 de Julio de 2013). *Krealo*. Recuperado el 12 de Diciembre de 2017, de *Cómo funciona una lámpara incandescente*: <https://www.krealo.es/blog/como-funciona-una-lampara-incandescente>

Ipi tel uva. (s.f.). *Umbralización (Thresholding)*. Recuperado el 24 de Octubre de 2018, de [ipi.tel.uva.es: https://www.ipi.tel.uva.es/~nacho/docencia/ing_ond_1/trabajos_03_04/sonificacion/cabroa_archivos/umbralizacion.html](https://www.ipi.tel.uva.es/~nacho/docencia/ing_ond_1/trabajos_03_04/sonificacion/cabroa_archivos/umbralizacion.html)

Marín Alvarez, L. D. (6 de Agosto de 2016). *Medellín ya tiene buses padrones de última tecnología*. *El Mundo*. Recuperado el 22 de Mayo de 2018, de Elmundo.com:

http://www.elmundo.com/portal/noticias/movilidad/medellin_ya_tiene_buses__padrones_de_ultima_tecnologia.php#.WwSt3rGFTIV

MathWorks. (s.f.). *Aprendizaje Supervisado*. Recuperado el 16 de Octubre de 2018, de Mathworks.com: <https://la.mathworks.com/discovery/aprendizaje-supervisado.html>

MatLab. (s.f.). *MATLAB Examples*. Recuperado el 10 de Octubre de 2018, de mathworks.com: <https://www.mathworks.com/examples/image/community/35149-reducecolorpeppers>

Mecalux. (s.f.). *Sensores*. Recuperado el 20 de Mayo de 2018, de logismarket.com.mx: <https://www.logismarket.com.mx/sensores-automatizacion/1314533052-cp.html>

Metro de Medellín. (21 de Febrero de 2018). *ACCESIBILIDAD*. Recuperado el 22 de Mayo de 2018, de Metrodemedellin.gov.co: <https://www.metrodemedellin.gov.co/viajeconnosotros/accesibilidad>

Ministerio de Educación Nacional. (s.f.). *Mineduccion.gov.co*. Recuperado el 6 de Mayo de 2018, de El Gobierno Nacional anuncia subsidios para discapacitados en todo el país: <https://www.mineduccion.gov.co/cvn/1665/article-165926.html>

Ministerio de Salud y Protección Social. (s.f.). *Minsalud.gov.co*. Recuperado el 14 de Mayo de 2018, de ABECÉ de la Discapacidad: <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/PS/abece-de-la-discapacidad.pdf>

Muñoz Manso, R. (2014). *Sistema de visión artificial para la la deteccion y lectura de matrículas*. Valladolid: UNIVERSIDAD DE VALLADOLID. Recuperado el 17 de Mayo de 2018

Musso, C. (7 de Febrero de 2018). *¿Qué Es Y Para Qué Sirve El Espacio De Color?* Recuperado el 14 de Octubre de 2018, de blogdelfotografo.com: <https://www.blogdelfotografo.com/espacio-color/>

News-Medical. (17 de Junio de 2012). *news-medical.net*. Recuperado el 5 de Mayo de 2018, de Tipos de discapacidad visual: [https://www.news-medical.net/health/Types-of-visual-impairment-\(Spanish\).aspx](https://www.news-medical.net/health/Types-of-visual-impairment-(Spanish).aspx)

Ng, A. (s.f.). *Deep Learning*. Recuperado el 22 de Octubre de 2018, de coursera.org: <https://www.coursera.org/?authMode=login>

Observatorio de la Accesibilidad. (s.f.). *LA ACCESIBILIDAD PARA LAS PERSONAS CON DISCAPACIDAD VISUAL: ROTULACION Y SEÑALETICA*. Recuperado el 05 de 11 de 2018, de observatoriodelaaccesibilidad.es: <https://www.observatoriodelaaccesibilidad.es/espacio-divulgativo/articulos/la-accesibilidad-personas-discapacidad-visual-rotulacion-senaletica.html>





OMS. (Agosto de 2014). *Organizacion mundial de la salud*. Recuperado el 15 de Diciembre de 2017, de <http://who.int/features/factfiles/blindness/es/>

- OMS. (Octubre de 2017). *Organización Mundial de la salud*. Recuperado el 15 de Diciembre de 2017, de Ceguera y discapacidad visual: <http://www.who.int/mediacentre/factsheets/fs282/es/>
- OpenCV. (18 de Diciembre de 2015). *Opencv.org*. Recuperado el 20 de Mayo de 2018, de Histograms - 1 : Find, Plot, Analyze !!!: https://docs.opencv.org/3.1.0/d1/db7/tutorial_py_histogram_begins.html
- OpenCV. (s.f.). *Acerca de OpenCV*. Obtenido de <https://opencv.org/about.html>: <https://opencv.org/about.html>
- Oxford. (s.f.). *Oxforddictionaries.com*. Recuperado el 5 de Mayo de 2018, de <https://es.oxforddictionaries.com/definicion/billete>
- Palazzesi, A. (14 de Junio de 2006). *Neoteo.com*. Recuperado el 2 de Noviembre de 2017, de CCD vs. CMOS: <https://www.neoteo.com/ccd-vs-cmos/>
- Presidencia de la República. (27 de Febrero de 2013). Ley Estatutaria N° 1618 de 27 febrero 2013. Santafé de Bogotá, Cundinamarca, Colombia. Recuperado el 13 de Mayo de 2018, de <http://wsp.presidencia.gov.co/Normativa/Leyes/Documents/2013/LEY%201618%20DEL%2027%20DE%20FEBRERO%20DE%202013.pdf>
- Quijada, C. C. (12 de Noviembre de 2012). *Tipos de luz artificial*. Recuperado el 18 de Diciembre de 2017, de <http://tucasanueva.com.mx/hogar/los-tipos-de-luz-artificial/>
- Raj, B. (s.f.). Data Augmentation | How to use Deep Learning when you have Limited Data— Part 2. Recuperado el 22 de Octubre de 2018, de [medium.com](https://medium.com/nanonets/how-to-use-deep-learning-when-you-have-limited-data-part-2-data-augmentation-c26971dc8ced): <https://medium.com/nanonets/how-to-use-deep-learning-when-you-have-limited-data-part-2-data-augmentation-c26971dc8ced>
- Ramos Fandiño, G. P. (s.f.). *Servicios bibliotecarios para ciegos y débiles visuales*. Recuperado el 5 de Octubre de 2018, de <http://132.248.9.34/hevila/Revistamexicanadecienciasdelainformacion/2010/vol1/no2/2.pdf>
- Raona. (31 de Mayo de 2017). *Los 10 Algoritmos esenciales en Machine Learning*. Recuperado el 15 de Octubre de 2018, de [Raona.com](https://www.raona.com/los-10-algoritmos-esenciales-machine-learning/): <https://www.raona.com/los-10-algoritmos-esenciales-machine-learning/>
- RASPBERRY PI FOUNDATION. (Febrero de 2015). *RASPBERRY PI 2 MODEL B*. Obtenido de <https://www.raspberrypi.org/products/raspberry-pi-2-model-b/>
- RCN. (28 de Noviembre de 2017). *rcnradio.com*. Recuperado el 6 de Mayo de 2018, de Desempleo de personas con discapacidad visual es siete veces el del país, según estudio: <https://www.rcnradio.com/economia/desempleo-de-personas-con-discapacidad-visual-es-siete-veces-el-del-pais-segun-estudio>

- Robologs. (22 de Enero de 2017). *Tutorial de Redes Neuronales con VREP C++ y Linux*. Recuperado el 22 de Octubre de 2018, de robologs.net: <https://robologs.net/2017/01/22/tutorial-de-redes-neuronales-con-vrep-c-y-linux/>
- Rodríguez Bazaga, A. (18 de Agosto de 2015). *¿Qué es OpenCV?* Obtenido de osl.ull.es: <https://osl.ull.es/software-libre/opencv-libreria-vision-computador/>
- Rosebrock, A. (28 de Agosto de 2017). *Fast, optimized 'for' pixel loops with OpenCV and Python*. Recuperado el 20 de Mayo de 2018, de Pyimagesearch.com: <https://www.pyimagesearch.com/2017/08/28/fast-optimized-for-pixel-loops-with-opencv-and-python/>
- scikit-learn. (s.f.). *scikit-learn Machine Learning in Python*. Recuperado el 24 de Octubre de 2018, de scikit-learn.org: <http://scikit-learn.org/stable/>
- Tensorflow. (28 de Septiembre de 2018). *Graphs and Sessions*. Obtenido de Tensorflow.org: <https://www.tensorflow.org/guide/graphs>
- Terris, R. (26 de Mayo de 2014). *Albedomedia.com*. Recuperado el 19 de Mayo de 2018, de Sensores de imagen – CMOS contra CCD: <https://www.albedomedia.com/tecnologia/sensores-de-imagen-cmos-contra-ccd/>
- Unisci24. (2014). *Unisci24.com*. Recuperado el 4 de Diciembre de 2017, de LASER 272920 79.44 KB: <http://unisci24.com/272920.html>
- Universia. (s.f.). *¿Qué es y para que sirve Python?* Recuperado el 24 de Mayo de 2018, de Universia.es: <http://noticias.universia.es/ciencia-tecnologia/noticia/2017/07/19/1154393/sirve-phyton.html>
- Universidad de Antioquia. (s.f.). *Servicio para personas con discapacidad visual*. Recuperado el 22 de Mayo de 2018, de udea.edu.co: http://www.udea.edu.co/wps/portal/udea/web/generales/interna!/ut/p/z0/fY-9CsJAEIRfJRaW4WUf8oQRBArBdFrZHO36mpyG2_P4ON7KhY2dvOxM7OM0mqntlOOThCIHdSR93p8mM7KfFCMstVivSyzYlwW88lmu8qnuVq_d8QG-hyu-ICacMu4COoXcs-QH23CP0M5JfO3OBHC0nABhKLSUVVTRzQgPQzDp4IEfQdGYoq3
- Universidad Politécnica de Madrid. (s.f.). *Interpretacion automática de las imagenes*. Recuperado el 20 de Mayo de 2018, de <http://www.elai.upm.es/moodle/mod/resource/view.php?id=971>
- Universidad Politécnica de Madrid. (s.f.). *Segmentación*. Recuperado el 14 de Mayo de 2018, de Elai.upm.es: <http://www.elai.upm.es/moodle/mod/resource/view.php?id=143>
- Universidad Politécnica de Madrid. (s.f.). *Técnicas de preprocesado*. Recuperado el 20 de Mayo de 2018, de <http://www.elai.upm.es/moodle/mod/resource/view.php?id=381>

- Valencia Armijos, Á., & Idrovo Urgilés, R. (2016). *DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA DE RECONOCIMIENTO Y MANIPULACIÓN DE FRUTAS UTILIZANDO VISIÓN ARTIFICIAL Y BRAZO ROBÓTICO INDUSTRIAL*. GUAYAQUIL – ECUADOR: ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL. Recuperado el 19 de Mayo de 2018
- Vasco Orozco, A., & Tamayo Ocampo, A. M. (2016). Visión artificial, una alternativa a la sensorica convencional en la. *IngEam*. Recuperado el 14 de Mayo de 2018, de <http://www.eam.edu.co/ojs/index.php/ingeam/article/download/144/207>
- Yneeds. (s.f.). *OCR de DNIs con Redes Neuronales Convolucionales*. Recuperado el 22 de Octubre de 2018, de yneeds.com: <http://yneeds.com/ocr-de-dnis-con-redes-neuronales-convolucionales/>
- Yu, D., Wang, H., & Wei, Z. (2014). *Mixed Pooling for Convolutional Neural Networks*. Recuperado el 22 de Octubre de 2018, de [semanticscholar.org: https://www.semanticscholar.org/paper/Mixed-Pooling-for-Convolutional-Neural-Networks-Yu-Wang/de664f22dd4c7b4c15ac4a52513004aee55765ff/figure/2](https://www.semanticscholar.org/paper/Mixed-Pooling-for-Convolutional-Neural-Networks-Yu-Wang/de664f22dd4c7b4c15ac4a52513004aee55765ff/figure/2)
- Zambrano, J. (s.f.). ¿Aprendizaje supervisado o no supervisado? Conoce sus diferencias dentro del machine learning y la automatización inteligente. *Medium*. Recuperado el 16 de Octubre de 2018, de <https://medium.com/@juanzambrano/aprendizaje-supervisado-o-no-supervisado-39ccf1fd6e7b>
- Zuñiga, R. (s.f.). *Introducción al Uso de MATLAB*. Recuperado el 24 de Mayo de 2018, de [geociencias.unam.mx: http://www.geociencias.unam.mx/~bole/eboletin/Matlabintro0408.pdf](http://www.geociencias.unam.mx/~bole/eboletin/Matlabintro0408.pdf)

Anexo 1. Matriz morfológica

<p>Adquirir imagen</p>	<p>A</p>	<p>B</p>	<p>C</p>
	 <p>Webcam mini: Son webcam que tienen un tamaño similar al de una moneda, son más fáciles de llevar y menos incómodas. (ElecFreaks, 2013)</p>	 <p>Gafas con cámara: Gafas que tienen integrada una cámara muy pequeña, sirven más que todo para hacer espionaje (Clarín, 2017)</p>	 <p>Web Cam: es una cámara de video económica y fácil de adaptar al computador. (Mas adelante, s.f.)</p>
<p>Procesar imagen</p>	<p>C</p>	<p>B</p>	<p>A</p>
	 <p>Odroid: Mini-ordenador compatible con linux y android, suele ser mas potente que la Raspberry Pi siendo igual de pequeña. (Velasco, 2015)</p>	 <p>Computador portátil: equipo personal mucho más grande que los demás, pero alcanza a procesar archivos robustos, además de que tiene todo integrado, como cámara. (Pérez Porto & Merino, 2013)</p>	 <p>Raspberry pi: Es una placa computadora de bajo costo, usualmente se utiliza para aplicaciones estudiantiles. (Universidad politecnica de Valencia, 2013)</p>
	<p>A</p>	<p>B</p>	<p>C</p>

<p>Transformar energía</p>	 <p>Batería Ion de Litio: Es un dispositivo diseñado para almacenamiento de energía, que usa sal de litio como electrolito y así generar una reacción electroquímica reversible para obtener un ánodo y un cátodo. (Ecured, s.f.)</p>	 <p>Baterías níquel hidruro: Son baterías que poseen una alta densidad de energía, lo cual tienen un tiempo de vida mayor. (Ecured, s.f.)</p>	 <p>Batería de lipo: Son baterías de polímero litio que poseen una excelente relación, calidad / precio, por eso son más usadas en aplicaciones con motores. (Erle Robotics, s.f.)</p>
<p>Extraer patrones y predecir imagen</p>	<p style="text-align: center;">C</p>  <p>C++: Lenguaje de programación orientado a objetos de bajo nivel proveniente de C. (Zator, s.f.)</p>	<p style="text-align: center;">B</p>  <p>Matlab: Es un sistema de programación y cálculo basado en la manipulación de matrices (Zuñiga)</p>	<p style="text-align: center;">A</p>  <p>Python: Lenguaje de programación orientado a objetos de alto nivel, interpretado y multipropósito. (CodeJobs, 2013)</p>
	<p style="text-align: center;">A</p>	<p style="text-align: center;">C</p>	<p style="text-align: center;">B</p>

<p>Enviar mensaje</p>	 <p>Auriculares: dispositivo que se utiliza para escuchar audio desde un dispositivo electrónico (Alegsa, s.f.)</p>	 <p>Auriculares con bluetooth: Auriculares inalámbricos que utilizan la tecnología bluetooth ya que permite transmitir grandes cantidades de datos a gran velocidad (Tecnología informatica, s.f.)</p>	 <p>Auriculares con radiofrecuencia: Auriculares inalámbricos que utilizan frecuencias de radio para transmitir la información auditiva, estas frecuencias pueden transmitir hasta 100 metros de distancia en un arco de 360°. (Informatica hoy, s.f.)</p>
------------------------------	---	--	--

