

Jorge Agustín
Velázquez Villegas

Desarrollo en hardware-software de un prototipo
traductor de caracteres impresos a sistema Braille

2024



Universidad Autónoma de Querétaro

Facultad de Ingeniería

Maestría en Mecatrónica

**Desarrollo en hardware-software de un prototipo
traductor de caracteres impresos a sistema Braille**

Que como parte de los requisitos para obtener el grado de
Maestro en Mecatrónica

Presenta:

Jorge Agustín Velázquez Villegas

Dirigido por:

Dr. Miguel Trejo Hernández

San Juan del Río, Qro. Junio del 2024

La presente obra está bajo la licencia:
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>



CC BY-NC-ND 4.0 DEED

Atribución-NoComercial-SinDerivadas 4.0 Internacional

Usted es libre de:

Compartir — copiar y redistribuir el material en cualquier medio o formato

La licenciante no puede revocar estas libertades en tanto usted siga los términos de la licencia

Bajo los siguientes términos:



Atribución — Usted debe dar [crédito de manera adecuada](#), brindar un enlace a la licencia, e [indicar si se han realizado cambios](#). Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante.



NoComercial — Usted no puede hacer uso del material con [propósitos comerciales](#).



SinDerivadas — Si [remezcla, transforma o crea a partir](#) del material, no podrá distribuir el material modificado.

No hay restricciones adicionales — No puede aplicar términos legales ni [medidas tecnológicas](#) que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia.

Avisos:

No tiene que cumplir con la licencia para elementos del material en el dominio público o cuando su uso esté permitido por una [excepción o limitación](#) aplicable.

No se dan garantías. La licencia podría no darle todos los permisos que necesita para el uso que tenga previsto. Por ejemplo, otros derechos como [publicidad, privacidad, o derechos morales](#) pueden limitar la forma en que utilice el material.



Universidad Autónoma de Querétaro
Facultad de ingeniería
Maestría en mecatrónica



Desarrollo en hardware-software de un prototipo traductor de caracteres impresos a sistema Braille

TESIS

Como parte de los requisitos para obtener el Grado de
Maestría en Mecatrónica

Presenta:

Jorge Agustín Velázquez Villegas

Dirigido por:

Dr. Miguel Trejo Hernández

SINODALES:

Dr. Miguel Trejo Hernández
Presidente

Dr. Irving Armando Cruz Albarrán
Secretario

Dr. Roque Alfredo Osornio Ríos
Vocal

Dr. Martín Valtierra Rodríguez
Suplente

Dr. J. Jesús de Santiago Pérez
Suplente

Campus San Juan del Río
San Juan del Río, Qro.
junio del 2024

RESUMEN

El braille es un sistema de lectura y escritura táctil que utilizan las personas con discapacidad visual, este les permite acceder a información escrita de esta manera; pero los documentos, artículos y diferentes medios impresos en tinta que no tienen algún texto escrito en braille, no pueden ser usados por personas con discapacidad visual, por lo que necesitan la ayuda de otra gente con la capacidad de ver y leer. El crear un dispositivo que pueda reconocer la información impresa en este tipo de documentos, artículos y medios escritos, resultaría útil para las personas con discapacidad visual. El trabajo escrito en este documento consiste en la creación de un sistema traductor de texto impreso a braille. El contenido que se encuentra escrito en este documento abarca en el primer capítulo: la historia del sistema braille, los trabajos previos que están relacionados con el tema de conversión de un medio comunicación a el sistema braille, el problema a resolver, la hipótesis propuesta como solución, los objetivos y el planteamiento general de este trabajo. El siguiente capítulo es la fundamentación teórica, en el que se encuentran los conceptos, mecanismos y procesos utilizados en el diseño y construcción del sistema traductor de texto a braille, además de explicar los distintos aparatos usados en este sistema; entre los dispositivos electrónicos utilizados están: una cámara digital, una computadora, una tarjeta con un microcontrolador y varios componentes electrónicos. Después de la teoría, está el capítulo donde se explica la metodología del sistema de traducción de texto a braille, las diferentes partes en las que se compone y la construcción de cada una; estas últimas son tres etapas: la primera es adquisición de imágenes; la segunda es el procesamiento de las imágenes, detección de caracteres y conversión a braille; y la tercera etapa es el control de un aparato electrónico, el cual crea forma física un signo braille. En el penúltimo capítulo están la explicación de las pruebas que se realizaron, y los resultados que arroja el sistema de traducción de texto a braille. Por último, está la conclusión del trabajo, los problemas que presenta el sistema de traducción y las mejoras que se pueden hacer a este.

Palabras clave: braille, procesamiento de imágenes, reconocimiento óptico de caracteres, sistema de traducción.

ABSTRACT

Braille is a tactile reading and writing system used by visually impaired people. It allows them to access written information independently. However, documents, articles, and various media printed in ink, which do not have text written in Braille, cannot be used by people with visual disabilities. As a result, they often need assistance from others who can see and read. This document presents the creation of a printed text-to-Braille translation system. The content includes the first chapter, which covers the history of the Braille system, previous works related to converting communication media to the Braille system, the problem to be solved, the hypothesis proposed as a solution, the objectives, and the general approach of this work. The next chapter provides a theoretical foundation, detailing the concepts, mechanisms, and processes used in the design and construction of the text-to-Braille translator system. It also explains the different devices used in the system. Following the theory, there is a chapter that outlines the methodology of the text-to-Braille translation system, describing its distinct components and the construction of each. These components are divided into three stages: the first is image acquisition, the second is image processing, character detection, and conversion to Braille, and the third stage involves controlling an electronic device to create a Braille sign in physical form. The penultimate chapter explains the tests performed and the results produced by the text-to-Braille translation system. Finally, the document concludes with an overview of the work, discussing the problems encountered with the translation system and potential improvements that can be made.

Keywords: braille, imaging processing, character optical recognition, translation system.

AGRADECIMIENTOS

En las siguientes palabras agradezco a todas las personas que estuvieron involucradas en el desarrollo de este trabajo, y que su ayuda me permitió realizar este proyecto. En primero está el Dr. Miguel Trejo Hernández, quien es el director, gracias a su continuo seguimiento y supervisión logré completar este trabajo; otra persona es el Dr. Irving Armando Cruz Albarrán, quién me ayudó en la etapa de procesamiento de imágenes; alguien más es el Dr. Roque Alfredo Osornio Ríos, que con sus consejos pude delimitar las ideas y el alcance de este proyecto, para que yo no tuviera que hacer más de lo debido.

También quiero agradecer a todos los sinodales por las observaciones que me hicieron, y que me ayudaron a terminar este trabajo; además agradezco a mi familia, mis padres, mis hermanos y hermana, quienes son un gran apoyo. Por último, un gracias a quienes se tomen la molestia de leer este documento.

ÍNDICE

RESUMEN.....	ii
ABSTRACT	iii
AGRADECIMIENTOS	iv
ÍNDICE	v
ÍNDICE DE CUADROS	vii
ÍNDICE DE FIGURAS	vii
1 INTRODUCCIÓN.....	1
1.1 Antecedentes	4
1.2 Descripción del problema	7
1.3 Justificación.....	8
1.4 Hipótesis	9
1.5 Objetivos	9
1.6 Planteamiento general	10
2 FUNDAMENTACIÓN TEÓRICA.....	12
2.1 Sistema braille.....	13
2.2 Redes neuronales artificiales.....	15
2.3 Procesamiento de imágenes digitales.....	17
2.4 Reconocimiento Óptico de Caracteres	22
2.5 Manufactura Aditiva	23
2.6 Microcontroladores	25
2.7 Cámaras digitales	27
2.8 Actuadores solenoides.....	28
2.9 Software e interfaz	29
3 METODOLOGÍA	31
3.1 Diseño general del sistema.....	32
3.2 Metodología del sistema	33
3.3 Dispositivo de adquisición.....	36
3.3.1 Cámara digital	36

3.3.2	Control de iluminación.....	37
3.3.3	Base para el módulo de la cámara	38
3.3.4	Pruebas del dispositivo de adquisición	39
3.4	Procesamiento de imágenes	41
3.4.1	Métodos de procesamiento de imágenes	41
3.4.2	Programa de computadora	45
3.4.3	Pruebas del procesamiento de imágenes.....	48
3.5	Dispositivo de salida	56
3.5.1	Comunicación y control	56
3.5.2	Etapa de potencia	57
3.5.3	Representación del signo braille	57
3.5.4	Pruebas del dispositivo de salida.....	58
4	PRUEBAS Y RESULTADOS	60
4.1	Pruebas del sistema	61
4.2	Resultados del sistema.....	64
5	CONCLUSIÓN.....	71
5.1	Prospectivas del trabajo.....	74
6	BIBLIOGRAFÍA.....	75
7	APÉNDICES	80
7.1	Apéndice A: Artículo del dispositivo de salida	81
7.2	Apéndice B: Texto de pruebas en Arial.....	88
7.3	Apéndice C: Texto de pruebas en Times New Roman	89

ÍNDICE DE CUADROS

Tabla 1. Matriz de pruebas del sistema de traducción de texto impreso a braille.....	62
---	----

ÍNDICE DE FIGURAS

Figura 1-1. Diagrama del planteamiento general (Autoría propia, 2024).	11
Figura 2-1. Matriz de seis puntos, distribuidos en dos columnas y tres renglones (Martínez Calvo, 2006).....	13
Figura 2-2. Disposición de los puntos para una celda braille (Martínez Calvo, 2006).....	13
Figura 2-3. Normas de la CBE para signos braille: (a) dimensiones para una celda braille, (b) los puntos que la conforman (García Soria & Ruiz Prieto, 2014).	14
Figura 2-4. Signos del braille español: (a) alfabeto braille, (b) números en braille, (c) signos diacríticos del español en braille, (d) ejemplos de representación de mayúsculas, (e) signos de puntuación en braille (García Soria & Ruiz Prieto, 2023).	15
Figura 2-5. Elementos de una red neuronal artificial: (a) neurona simple, (b) una función de salida sigmoideal (Rao, 2009).....	16
Figura 2-6. Red retroalimentada de multicapa (Rao, 2009).	16
Figura 2-7. Pasos fundamentales en el procesamiento de imágenes digitales (Gonzalez & Woods, 2018).....	18
Figura 2-8. Representación de imágenes digitales: (a) imagen presentada como una matriz de intensidad visual, (b) imagen mostrada como una matriz numérica (Gonzalez & Woods, 2018).....	19
Figura 2-9. Sistema OCR típico (Borovikov, 2004).	23
Figura 2-10. Esquema de un sistema basado en extrusión (Gibson et al., 2021).....	24
Figura 2-11. Esquema general de un microcomputador (Valdéz Pérez & Pallás Areny, 2007).....	25
Figura 2-12. Esquema de bloques general de un microcontrolador (Valdéz Pérez & Pallás Areny, 2007).....	26
Figura 2-13. Diseño de los sensores CDD y CMOS para una cámara digital (Griffing, 2023).....	28
Figura 2-14. Imágenes de solenoides: (a) bosquejo de un actuador solenoide, (b) solenoides industriales (Gomis-Bellmunt & Campanile, 2010).	29
Figura 3-1. Diagrama general del sistema de traducción (Autoría propia, 2024).	32
Figura 3-2. Diagrama metodológico simple (Autoría propia, 2024).	33
Figura 3-3. Diagrama metodológico extendido (Autoría propia, 2024).	34
Figura 3-4. Diagrama detallado de la metodología para el sistema de traducción de texto impreso en español a signos braille (Autoría propia, 2024).	35
Figura 3-5. Módulo de cámara HBVCAM-W202012HD V33 (Autoría propia, 2024).	36
Figura 3-6. Componentes de control de iluminación de la cámara: (a) fotorresistencia, (b) led de luz blanca (Autoría propia, 2024).	37
Figura 3-7. Módulo HBVCAM-W202012HD con auto iluminación led (Autoría propia, 2024).....	37

Figura 3-8. Partes del dispositivo de salida: (a) base para el módulo de la cámara, (b) tapa para el módulo de la cámara (Autoría propia, 2024).	38
Figura 3-9. Tapa de la fotorresistencia (Autoría propia, 2024).	38
Figura 3-10. Módulo de la cámara montada en la base con una tapa en la fotorresistencia: (a) vista inferior, (b) vista frontal (Autoría propia, 2024).	39
Figura 3-11. Módulo de la cámara en la base con la tapa conectado a la computadora, vista superior (Autoría propia, 2024).	39
Figura 3-12. Captura de la cámara digital con iluminación externa (Autoría propia, 2024).	40
Figura 3-13. Captura de la cámara digital con la auto iluminación (Autoría propia, 2024).	40
Figura 3-14. Imagen a color (Autoría propia, 2024).	42
Figura 3-15. Imagen en escala de grises (Autoría propia, 2024).	42
Figura 3-16. Imagen en escala de grises con el desenfoque gaussiano (autoría propia, 2023).	43
Figura 3-17. Imagen en blanco y negro (Autoría propia, 2024).	43
Figura 3-18. Imagen con la implementación del algoritmo Canny (Autoría propia, 2024).	44
Figura 3-19. Imagen con la implementación de algoritmo rectángulo delimitador (Bounding Rectangle) (Autoría propia, 2024).	44
Figura 3-20. Imagen a color rotada (Autoría propia, 2023).	45
Figura 3-21. Programa de computadora para el procesamiento de imágenes (Autoría propia, 2024).	46
Figura 3-22. Función de cada parte del programa de computadora para el procesamiento de imágenes (Autoría propia, 2024).	47
Figura 3-23. Captura de imagen a color con texto inclinado (Autoría propia, 2024).	49
Figura 3-24. Imagen en escala de grises con texto inclinado (Autoría propia, 2024).	49
Figura 3-25. Imagen con desenfoque gaussiano del texto inclinado (Autoría propia, 2024).	49
Figura 3-26. Imagen en blanco y negro del texto inclinado (Autoría propia, 2024).	49
Figura 3-27. Algoritmo Canny aplicado en la imagen del texto inclinado (Autoría propia, 2024).	50
Figura 3-28. Algoritmo del rectángulo delimitador implementado en la imagen del texto inclinado (Autoría propia, 2024).	50
Figura 3-29. Objetos de la imagen dentro de la zona de interés (Autoría propia, 2024).	50
Figura 3-30. Objeto de la imagen con la mayor área (Autoría propia, 2024).	51
Figura 3-31. Recuadros usados en el cálculo del ángulo de rotación (Autoría propia, 2024).	51
Figura 3-32. Imagen en blanco y negro rotada un cierto ángulo (Autoría propia, 2024).	52
Figura 3-33. Imagen en blanco y negro con texto horizontal (Autoría propia, 2024).	52
Figura 3-34. Algoritmo Canny aplicado a la imagen de blanco y negro con texto horizontal (Autoría propia, 2024).	53
Figura 3-35. Algoritmo de rectángulo delimitador implementado en la imagen con texto horizontal (Autoría propia, 2024).	53
Figura 3-36. Recuadros dentro de la imagen (Autoría propia, 2024).	53
Figura 3-37. Construcción de bordes más grandes (Autoría propia, 2024).	54
Figura 3-38. Recuadros más grandes de la imagen (Autoría propia, 2024).	54

Figura 3-39. Recuadro grande más centrado en la imagen (Autoría propia, 2024).....	54
Figura 3-40. Sección recortada de la imagen en blanco y negro con texto horizontal (Autoría propia, 2024).....	55
Figura 3-41. Tarjeta Arduino UNO (Autoría propia, 2024).....	56
Figura 3-42. Etapa de potencia de la celda braille (Autoría propia, 2023).	57
Figura 3-43. Celda braille (Autoría propia, 2024).....	58
Figura 3-44. Electroimán solenoide (Autoría propia, 2024).....	58
Figura 3-45. Dispositivos electrónicos y electromecánicos que componen la celda braille electrónica (Autoría propia, 2024).....	59
Figura 3-46. Conversión de caracteres de palabras (imagen izquierda) al sistema braille, y representados en la celda braille (imagen derecha) (Autoría propia, 2024).	59
Figura 4-1. Sistema de traducción de texto al sistema braille (Autoría propia, 2024).	61
Figura 4-2. Dispositivo de adquisición sobre una hoja con texto impreso (Autoría propia, 2024).	63
Figura 4-3. Programa de computadora realizando el procesamiento de imágenes (Autoría propia, 2024).....	64
Figura 4-4. Dispositivo de salida representando en la celda braille el signo transmitido por la computadora (Autoría propia, 2024).....	64
Figura 4-5. Resultados del Tesseract OCR con letras minúsculas en Arial (Autoría propia, 2024).	65
Figura 4-6. Resultados del Tesseract OCR con letras mayúsculas en Arial (Autoría propia, 2024).....	65
Figura 4-7. Resultado del Tesseract OCR con números Arial (Autoría propia, 2024).....	66
Figura 4-8. Resultados del Tesseract OCR con los signos de puntuación (Autoría propia, 2024).	66
Figura 4-9. Resultados del Tesseract OCR con letras minúsculas Times New Roman (Autoría propia, 2024).....	68
Figura 4-10. Resultados del Tesseract OCR con letras mayúsculas Times New Roman (Autoría propia, 2024).....	68
Figura 4-11. Resultados del Tesseract OCR con números Times New Roman (Autoría propia, 2024).....	69
Figura 4-12. Resultados del Tesseract OCR con signos de puntuación Times New Roman (Autoría propia, 2024).....	69

1 INTRODUCCIÓN

La información, y el acceso a esta misma, permite a las personas adquirir el conocimiento de manera más amplia o precisa sobre alguna materia en específico, además de contribuir en su desarrollo personal. Esto también se aplica en las personas con discapacidad visual, porque su facultad de razonamiento no está limitada por este tipo de discapacidad, además sus otros sentidos les ayudan a suplir diferentes necesidades. El sistema braille, de acuerdo con Navarro Saad (1998), es un método de lectura y escritura táctil de seis puntos creado en el siglo XIX con el objetivo de ayudar a las personas con discapacidad visual a aprender a leer y escribir; este sistema se encuentra escrito sobre una superficie plana, los patrones de puntos que sobresalen de ésta representan las letras, palabras, números, entre otros signos y vocablos. La forma de leer el braille es mediante el sentido del tacto, en particular por medio de las yemas en la punta de los dedos, porque son una parte del cuerpo con una gran modalidad sensorial.

La forma de escribir en braille de manera manual, acorde con González, (2016), es por medio de la “pauta braille”, es una plancha que consta de dos placas: una que contiene filas de surcos en forma de celda braille, la segunda tiene filas de agujeros cuadrados del tamaño de las celdas braille; el papel se coloca entre ambas placas y mediante un punzón especial se graba el signo braille en cada agujero cuadrado. González (2016) también mencionó que existen otro tipo instrumentos dedicados para escribir braille de una forma más rápida, uno de ellos es la máquina de Perkins, creada por David Abraham en 1951, este es un aparato que permite a una persona ciega escribir en braille, similar a una máquina de escribir, pero con menos teclas: seis teclas para cada punto del signo braille, un espaciador, una tecla para el salto de línea y una para el retroceso. El diseño y distribución de las teclas en la máquina Perkins es utilizado en diferentes dispositivos electrónicos dedicados a la lectura y escritura en braille, estos mismos también cuentan con un arreglo de celdas braille electrónicas que hacen la de signos braille, de tal forma como si estuvieran impresos en un papel, además estos aparatos son más ligeros y portables, pero tiene un alto costo que no es asequible para la mayoría de las personas. En México de acuerdo con los datos del Instituto Nacional de Estadística y Geografía (INEGI, 2020), existen 6 179 890 personas con algún tipo de discapacidad, dentro de esta población, 2 691 338 se encuentran la gente que tienen deficiencia visual, pérdida total de la vista en uno o ambos ojos, o son personas que aún, usando lentes no pueden ver bien.

El problema que afronta una persona con discapacidad visual es el acceso a la información, en particular la que se encuentra en los libros, revistas o en cualquier otro tipo de material con texto impreso, porque, aunque si existe material escrito en braille: libros, documentos y algunos artículos de uso personal como los medicamentos tienen impresos signos braille; pero otra gran parte de los documentos y artículos no están impreso en braille o no tienen este sistema de lectura. Un aspecto por considerar es la gente con discapacidad visual en México, de acuerdo con el INEGI (2020), entre un total de más de seis millones de

personas con algún tipo de discapacidad aproximadamente el 43.55 % tiene discapacidad o deficiencia visual; por esto mismo los mexicanos que tiene esta discapacidad requerirían de un aparato que les ayude a identificar el texto en los libros, artículos de consumo o en cualquier otro medio con información impresa en tinta.

El trabajo que se desarrolla en este documento es la construcción de un sistema capaz de identificar texto impreso visualmente para después convertirlo y representarlo en braille, todo esto mediante el uso de tres dispositivos principales: el primero, una cámara que toma fotografías de páginas de hojas sueltas, libros o material con texto impreso; el segundo, una computadora que tiene la función de procesar las imágenes, esto con objeto de identificar el texto en las misma, almacenar la información obtenida y convertirla en braille; por último, un aparato electrónico capaz de representar físicamente una celda braille y que cambie entre los diferentes signos braille, esto para identificar de manera táctil las letras que estén en las hojas de los libros o material con texto impreso. En el procesamiento de las imágenes se implementa un motor de OCR (Reconocimiento Óptico de Caracteres, *Optical Character Recognition*) para identificar el texto, además de también utilizar Redes Neuronales Artificiales en conjunto con otros métodos de procesamiento de imágenes para reconocer los caracteres. En cuanto al proceso de conversión de texto al sistema braille, se utilizan las normas definidas por la CBE (Comisión Braille Española) para una correcta transcripción al braille de los textos escritos en el idioma español, esta comisión también determina las dimensiones que deben tener los signos braille y las formas en que deben ser impresos los signo braille, esto último dependiendo el tipo de material utilizado.

En el primer capítulo se presentan los trabajos que son precedentes al desarrollo que está escrito en este documento, la problemática a resolver, la hipótesis que se plantea como solución, así como los objetivos que se tienen para esta investigación y la forma en la que el trabajo es presentado. Existen varios trabajos que se han acercado a solucionar este tipo de problema, algunos optaron por identificar objetos y situaciones en las imágenes, para después insertar un título en braille que describe el contenido de una fotografía; otros se enfocaron en detectar los caracteres en las páginas y trasladar la información a braille, en estos trabajos algunos captan toda la página con el texto impreso y otros solamente secciones de esta misma, como ciertos renglones o esquinas; también existen proyectos que se orientaron en analizar grabaciones, para después detectar el lenguaje en estas mismas y trasladar la información al sistema braille. En cuanto al trabajo propuesto se enfoca en detectar continuamente imágenes de texto en pequeñas secciones, de tal forma como si estuviera leyendo sobre un renglón, y convertir cada carácter al signo braille correspondiente en donde se posicione la cámara; esto para facilitar la lectura del texto impreso en diferentes tipos de materiales, y con ello una persona con discapacidad visual pueda tener acceso a la información contenida en el material con texto impreso.

1.1 Antecedentes

El sistema braille es un método de escritura y lectura de seis puntos, es utilizado principalmente por las personas con discapacidad visual y por quienes también tienen deficiencia auditiva o personas sordociegas. La historia del braille inicia en el siglo XIX, Navarro Saad (1998) mencionó que la creación de este sistema es debido al aporte de tres personas: Valentín Haüy, un pedagogo francés que fundó un colegio dedicado a la enseñanza de las personas ciegas; Charles Barbier de la Serre, un capitán de la armada francesa el cual introdujo un método de escritura y lectura que transmitía mensajes e instrucciones en la oscuridad; el último personaje es Luis Braille, un pedagogo que quedó ciego de joven debido a un accidente, él fue estudiante del colegio de Haüy, mejoró el sistema ideado por de la Serre y lo convirtió en la forma de escritura y lectura de seis puntos que lleva su nombre. Luis Braille publicó su método en 1829, pero no fue aceptado oficialmente hasta 1854, dos años después de su muerte; actualmente esta forma de leer y escribir es utilizada alrededor del mundo y, dependiendo del país y del idioma en el que se hable, las normas de escritura y lectura braille son diferentes.

Las personas con discapacidad visual tienen a su alcance libros con texto impreso en braille, pero aún existe una gran barrera que les impide acceder a información que les podría beneficiar, tal como reconocer el texto que tiene un objeto que estén sosteniendo, ya sea un libro que no tenga traducción en braille, revistas, empaques para artículos de alimentación o de medicamentos, entre otros. Actualmente se han desarrollado trabajos que tratan de ayudar a la gente con discapacidad visual, para que tengan un acceso a la información que proviene de diferentes medios. En investigaciones relacionadas con el idioma español uno es el trabajo de Acevedo et al. (2016), en el cual crearon un sistema que traduce texto escaneado a braille y viceversa; también está el trabajo desarrollado por Ramos-García et al. (2022), en este utilizaron IoT (Internet de las Cosas, *Internet Of Things*) para enviar la información desde un dispositivo inteligente hacia un aparato, en donde el texto se representa por medio de una celda braille.

En inglés existen una gran cantidad de investigaciones con enfoques muy diferentes, en unas el objetivo principal es obtener características de imágenes y traducir estas mismas a texto en braille; tal es el caso del trabajo de Hassan et al. (2019), ellos crearon un sistema que captura imágenes y por medio de una red CNN (Red Neuronal Convolutiva, *Convolutional Neural Network*) detectaron nueve objetos que traduce a braille, también usaron un motor OCR como apoyo para identificar texto en las fotografías, este último lo utilizaron como segundo método de reconocimiento. Zaman et al. (2019) crearon un sistema que genera títulos en las imágenes dependiendo de la situación mostrada, también usaron una red CNN para reconocer los objetos, pero además implementaron una red RNN (Red Neuronal Recurrente, *Recurrent Neural Network*) para generar el texto, este último también es traducido a braille. Jothi et al. (2022) hicieron un trabajo como

los dos anteriores, con una red CNN, una red RNN y un motor OCR, además de que estos sistemas trabajan en conjunto: detectan situaciones en las imágenes, escriben un subtítulo en braille y reconocen el texto de las imágenes, el cual también es convertido a braille.

Continuando con el idioma inglés está el trabajo de Liu et al. (2016), donde tomaron fotografías a textos, después realizaron correcciones de distorsión en las imágenes, usaron un motor OCR como método de identificación de caracteres y por último los convirtieron a signos braille. La investigación de Hossain et al. (2018) es similar a la anterior, pero las imágenes fueron capturadas por la cámara de un teléfono inteligente, en este último una aplicación asiste a la persona ciega a tomar las fotografías, después el texto en las imágenes es reconocido por un OCR y al final es convertido al sistema braille. En el trabajo de Kumari et al. (2020) también hicieron capturas de imágenes mediante una cámara, pero a diferencia de las dos investigaciones previas, esta no tiene algún proceso de corrección para las fotografías o una aplicación que asista al momento de tomar la foto, simplemente convirtieron a braille el texto que es detectado por un OCR. El desarrollo de Zamir et al. (2020) es parecido a las tres investigaciones anteriores, tomaron capturas de imágenes por medio de una cámara, usaron un OCR para identificar el texto, pero al final este es convertido solamente a audio. Rahimi et al. (2022) desarrollaron un dispositivo que convierte a braille palabras escritas en inglés de líneas de texto impreso, el aparato está continuamente realizando capturas de imágenes; este proyecto es un dispositivo colocado la punta del dedo y, cuya celda braille va en él, al igual que los cuatro trabajos anteriores implementaron un motor OCR como forma de reconocimiento de texto.

Entre otros trabajos relacionados con idioma inglés está el desarrollado por Başıftçi y Eldem (2016), en esta investigación escanearon hojas de texto, después reconocieron los caracteres mediante un motor OCR y, por último, en un dispositivo con 96 celdas táctiles braille representaron la información escaneada; también usaron la salida de audio como un método alternativo. En el trabajo de Adnan et al. (2017) simplemente almacenaron archivos de texto en una memoria, después un aparato lee la información de la memoria y la representa en una celda táctil braille. La investigación de Sultana et al. (2017) es similar al trabajo anterior, pero en el aparato implementaron 12 celdas táctiles braille, además que en este mismo se puede escribir en braille por medio de seis teclas que representan los puntos de un signo braille. Karmel et al. (2019) también construyeron un aparato que usa Vision Api de Google para leer la información en una imagen, el texto que es reconocido es convertido a un audio.

En lenguaje chino mandarín se realizaron investigaciones para comprobar el desempeño de métodos de identificación de texto, y con esto después convertir los caracteres chinos a braille. Li et al. (2019) llevaron a cabo una investigación sobre el modelo de Márkov, en ella verificaron la transformación al sistema braille de archivos digitales con texto en chino, con los resultados que obtuvieron

concluyeron que el modelo funciona para la conversión. En la investigación de Lo y Chou (2022) realizaron pruebas en una red neuronal para identificar caracteres chinos silábicos, que son los símbolos que definen la forma de cómo se debe leer en voz alta, pero el método que usaron solamente identificó los símbolos que representan nombres, objetos, entre otros, pero no los silábicos.

Kobayashi et al. (2022) llevaron a cabo una investigación para evaluar una red neuronal que transcribe audios en japonés a caracteres braille, esta red está construida de la manera E2E (Extremo a extremo, *End to end*) y con la ayuda del programa S2B (Voz a Braille, *Speech to Braille*) transformó las palabras reconocidas de los audios a braille. Por otra parte, pero relacionado con la voz, en el trabajo de Hussain et al. (2022) convirtieron voces del bengalí al braille, implementaron una red CNN en la cual tiene como base las voces de personas con diferentes edades. El trabajo que desarrollo Devi (2016) fue un dispositivo capaz de traducir de imágenes el lenguaje tamil a sistema braille y viceversa, además este aparato contó con una salida en audio del texto identificado en las fotografías, la forma de detección de texto fue a través de la segmentación de los renglones ya fuera en tamil o braille.

Los trabajos de investigación, anteriormente mencionados, abordaron el tema de convertir los textos, imágenes y archivos con información, en una forma en que las personas ciegas sean capaces de percibir, e incluso en otros transformaron grabaciones de voz al sistema braille. En los proyectos se implementaron redes neuronales como una forma de obtener las características en una imagen o de un audio, muchos trabajos implementaron un motor OCR para reconocer el texto en fotografías, y en pocos se usó la segmentación en las imágenes como un método de procesamiento. Una gran parte de estos desarrollos están centrados en el idioma inglés, unos cuantos en español y otros lenguajes. Por ello la importancia de crear un sistema capaz de generar código braille en español, esto a partir de imágenes capturadas en pequeñas secciones del texto, y que estas últimas se realicen de manera continua para que la cámara pueda moverse sobre la superficie donde este impresa la información; también que las capturas estén apoyadas por un control de iluminación que permita tomar las fotografías en un ambiente oscuro, ya que esta característica no se encuentra en otros trabajos reportados con enfoques similares al problema, además de utilizar otros métodos para el procesamiento de imágenes y el reconocimiento de texto en ellas. Por último, que el control de la celda braille electrónica este coordinado con las imágenes capturadas, para que los signos representados estén acordes con los caracteres en las fotografías.

En nuestro país, no se ha encontrado un dispositivo tendiente a ser portátil, a manera de accesorio en la mano del usuario, especialmente dedicado a la traducción de letras, palabras o texto en español impreso desde su adquisición, reconocimiento y transformación hacia el sistema braille en una matriz actualizable en el dedo índice derecho de algún usuario. En un país como lo es México, existe

una población de 6 197 890 personas con algún tipo de discapacidad de acuerdo con el censo generado por el INEGI en el 2020; dentro de esta cantidad de personas, aproximadamente el 43.55% se encuentra la gente que tiene la pérdida total de la vista en uno o ambos ojos, como también de aquellas personas con deficiencia visual y otras que no pueden ver bien aun usando lentes.

1.2 Descripción del problema

El acceso a la información ha sido una de las grandes herramientas que permiten a las personas desarrollarse como individuos. En el caso de las personas con discapacidad visual no es la excepción, porque aún sin la capacidad de ver o el tener una visión limitada, sus otros sentidos suplen diferentes tipos de necesidades, además de que tienen la facultad de razonar. Con esta idea en mente se creó el método de escritura y lectura braille para la gente con discapacidad visual, este sistema es ampliamente aceptado, pero no toda la información se encuentra disponible de esta manera, ya sea porque el acceso a este tipo de materiales de lectura se debe a que es escaso, o incluso por el hecho de tener un mayor costo que muchos libros de texto, dada la limitación de producirlos. También existen dispositivos con la capacidad de representar braille y poder escribir con ellos, pero suelen tener un alto costo o es complicado de conseguirlos para las personas de escasos recursos, impidiéndoles de esta manera aprender a leer o escribir braille, por tanto, sería muy conveniente contar con un dispositivo que ayude a leer texto impreso en tinta y transcribirlo a braille.

En los trabajos investigación optaron por desarrollos de dispositivos que procesan el texto en imágenes o archivos, este último es convertido a braille y mostrado en celdas braille electrónicas, en algunos de estos proyectos el procesamiento es de una captura de toda una página de texto, en otros solo una sección de esta, pero en pocos tienen la asistencia de iluminación para obtener una imagen más adecuada para ser procesada. Una gran parte de los trabajos utilizaron un motor OCR para reconocer texto, sin embargo, en algunos se realizó un preprocesamiento en la imagen para hacer correcciones en la misma, como corregir la distorsión, arreglar un mal ángulo de captura, e incluso hacer más definida la imagen, pero haciéndolo en gran parte para una imagen de toda una página; además la toma de la fotografía depende de donde este colocada la cámara, y cuando una persona con discapacidad visual necesite este sistema no siempre es seguro que la captura este bien hecha, y solo unos pocos desarrollos tiene un tipo de asistencia para las personas con algún problema visual. En cuanto al aparato donde se representan los signos braille, muchos implementaron una celda braille electrónica, pero dependiendo del tiempo en que debe ser mostrado el signo, y en pocos casos, esta celda es algo invasiva en las yemas de los dedos de la persona que la utiliza; otros usaron dispositivos con más celdas para mostrar

la información en braille, pero con un cierto peso considerable y necesitan de una constante alimentación de energía.

Debido a lo explicado en el párrafo anterior, el desarrollo de este trabajo se enfoca en crear un aparato que posea una asistencia en la iluminación de la cámara, las capturas se realizarán de manera continua en secciones del texto para que la cámara sea fácil de situar sobre este mismo; se utilizará una celda braille electrónica, la cual debe estar coordinada con las capturas, para tener una correcta representación del signo braille que corresponde; además, este sistema debe adecuarse a una velocidad de lectura de cada individuo, ser cómodo al momento de sostenerse y no invasivo con la persona que lo utiliza. El sistema usará un modelo computacional avanzado como los OCR y los algoritmos de Aprendizaje de Máquinas para su buen funcionamiento, identificando el texto dependiendo de donde esté posicionada la cámara que lo captura y el cambio de posición en referencia a este último. Algunos sistemas previamente desarrollados contemplan la solución de este problema, pero es deseable contar con uno que sea portable, que pueda reubicarse o desplegarse para dejar libre la palma de la mano si no se usa, que sea compacto el sistema de adquisición y se muestre la representación en sistema braille.

1.3 Justificación

El presente trabajo busca contribuir con el desarrollo de un prototipo en software-hardware para mejorar la forma en la que un usuario con deficiencia visual, o con deseo de aprendizaje del sistema braille, pueda tener acceso a información o conocimiento que este requiera, y que no se encuentre en este sistema de escritura de puntos en relieve. La falta de dispositivos en México, que faciliten el acceso a la información, para personas con alguna discapacidad visual, es muy evidente, dado que son pocos los lugares donde puede encontrarse este sistema de escritura grabado en alguna superficie. El sistema braille solo aparece en pocos lugares, además algunos productos lo traen en sus empaques, pero en una cantidad limitada. Por otra parte, se contribuirá elaborando un diseño fácil de manipular donde pueda contener una minicámara para la adquisición de imágenes del texto, en conjunto con la celda braille actualizable, esto con el fin de que el usuario no pierda la habilidad de lectura o tacto en el dedo. El sistema OCR en software con un algoritmo de procesamiento de redes neuronales, para el reconocimiento del texto escrito en español y su interacción con el hardware, se unirá con un sistema completo de traducción a braille para la comodidad del usuario. La generación del sistema se realizará a bajo costo y se considera que no sea invasivo.

Este trabajo está relacionado con el Programa Nacional Estratégico (PRONACES) en el área de Educación. Este proyecto pretende crear material y

tecnología que ayude a las personas con discapacidad visual en el proceso de aprendizaje de lectura, sobre todo en los medios que les son difíciles de acceder: libros, revistas y otro tipo de material con texto impreso.

1.4 Hipótesis

Mediante el uso de una cámara digital, algoritmos de reconocimiento de caracteres, uso de software, dispositivos electrónicos y de control, es posible desarrollar un sistema en hardware-software que permita la traducción de caracteres o palabras de texto impreso, alfabeto latino impreso desde un dispositivo especialmente dispuesto en el dedo índice, no invasivo, interconectado al sistema de traducción braille, siguiendo las normas de lectura y los estándares establecidos por la CBE, con el fin de que un usuario con discapacidad visual no requiera que la información se encuentre grabada en relieve previamente.

1.5 Objetivos

Objetivo general.

Diseñar y elaborar con manufactura aditiva un sistema hardware-software traductor de texto impreso en español a sistema braille, lo anterior con el uso de un OCR y redes neuronales en software, para reproducir las palabras en una celda táctil braille actualizable controlada por un microcontrolador, con el fin de disminuir la brecha de acceso al conocimiento e información en las personas con discapacidad visual.

Objetivos particulares

- a. Crear una interfaz de comunicación entre una computadora y una minicámara por medio de un lenguaje de programación para poder adquirir imágenes desde el dispositivo dispuesto en el dedo índice derecho de un usuario.
- b. Identificar la información contenida en las imágenes adquiridas en alfabeto latino con el uso de herramientas OCR y redes neuronales con el fin de extraer el texto.
- c. Entrenar una red neuronal artificial que identifique el texto para poder aplicar las normas de escritura de braille en español definidas por la CBE.
- d. Transcribir el texto al sistema braille mediante el uso de las normas definidas por la CBE para poder almacenar los caracteres que deben ser mostrados en una celda actualizable dispuesta en el dedo índice de un usuario, con el fin de que pueda identificarlo y no pierda tal habilidad.

- e. Programar y construir una tarjeta de control para un arreglo de seis solenoides mediante el uso de un microcontrolador para poder presentar al usuario la información.
- f. Fabricar una celda actualizable braille no invasiva, dispuesta para la presentación de un solo carácter, la cual debe respetar las medidas definidas por la CBE, lo anterior con el fin de que el usuario realice la lectura correspondiente.
- g. Realizar las pruebas de traducción y lectura necesarios para corroborar la efectividad del sistema hardware-software.

1.6 Planteamiento general

El planteamiento general del presente trabajo es mostrado en la Figura 1-1, el diagrama lo comprenden tres secciones: naranja, azul y verde. La primera parte se encuentra en color naranja, esta consiste en el estado del arte, es hacer la investigación de la teoría y los trabajos precedentes en base a la problemática observada y la solución propuesta. En el trabajo se identifica primeramente un problema, después se plantea una solución en forma de hipótesis, seguido se hace una búsqueda de investigaciones que abordaron el mismo tema, para con ello reconocer los trabajos que plantearon una solución similar a la propuesta en la hipótesis. Por último, hacer un resumen de cada trabajo, que métodos utilizaron, los resultados que obtuvieron y las conclusiones a las que llegaron, esto para determinar cómo encaja la solución propuesta en el tema de investigación, con que contribuye la hipótesis planteada, que aspectos de los trabajos anteriores pueden mejorarse, los diferentes métodos que no han sido implementado, entre otras cosas, y al final replantear la hipótesis propuesta y crear una metodología de esta misma. También en la misma primera sección, se expone realizar al mismo tiempo una investigación teórica relacionada con el tema: buscar conceptos, métodos, tecnologías, entre otras cosas, para después, con esto mismo tener una idea en concreto de lo que se puede crear; a esto le sigue el hecho de volver a definir la hipótesis en conjunto con los trabajos precedentes, además de también ayudar a la creación de la metodología.

En la sección azul, que es la segunda parte, se muestra de forma general la manera en cómo funciona la metodología. Con los aspectos teóricos y de antecedentes definidos se crea una metodología para la hipótesis, en esta parte se crean los pasos que debe tener el sistema que es planteado como solución; continuando, está el diseñar y construir los diferentes dispositivos y programas computacionales que se deben utilizar, para después evaluar el funcionamiento de estos mismos, y en caso de algún fallo o mal desempeño hacer un rediseño y reconstrucción de los mismos; después, está la unificación de todas las partes en un solo sistema y la comprobación del funcionamiento en conjunto de las

diferentes partes, es decir, corroborar que los dispositivos y los programas estén trabajando correctamente los unos con los otros, y al final, hacer distintas pruebas al sistema planteado como solución.

En la tercera sección del planteamiento general, primero, es recopilar y almacenar los datos obtenidos de las pruebas; después se realiza un análisis de estos mismos, qué aspectos son comparables con respecto a las investigaciones precedentes, determinar qué ventajas y desventajas tiene el sistema planteado como solución, las mejoras que pueden llevarse a cabo en el mismo, entre otras cosas. Al final, se encuentra la transcripción de los datos y el análisis de estos mismos, presentar en la tesis del proyecto de investigación las imágenes, tablas y una explicación de estas misma de manera escrita; qué resultados se obtuvieron en comparación con otros trabajos, las perspectivas del proyecto, entre otras cosas, para culminar con las conclusiones del trabajo desarrollado.

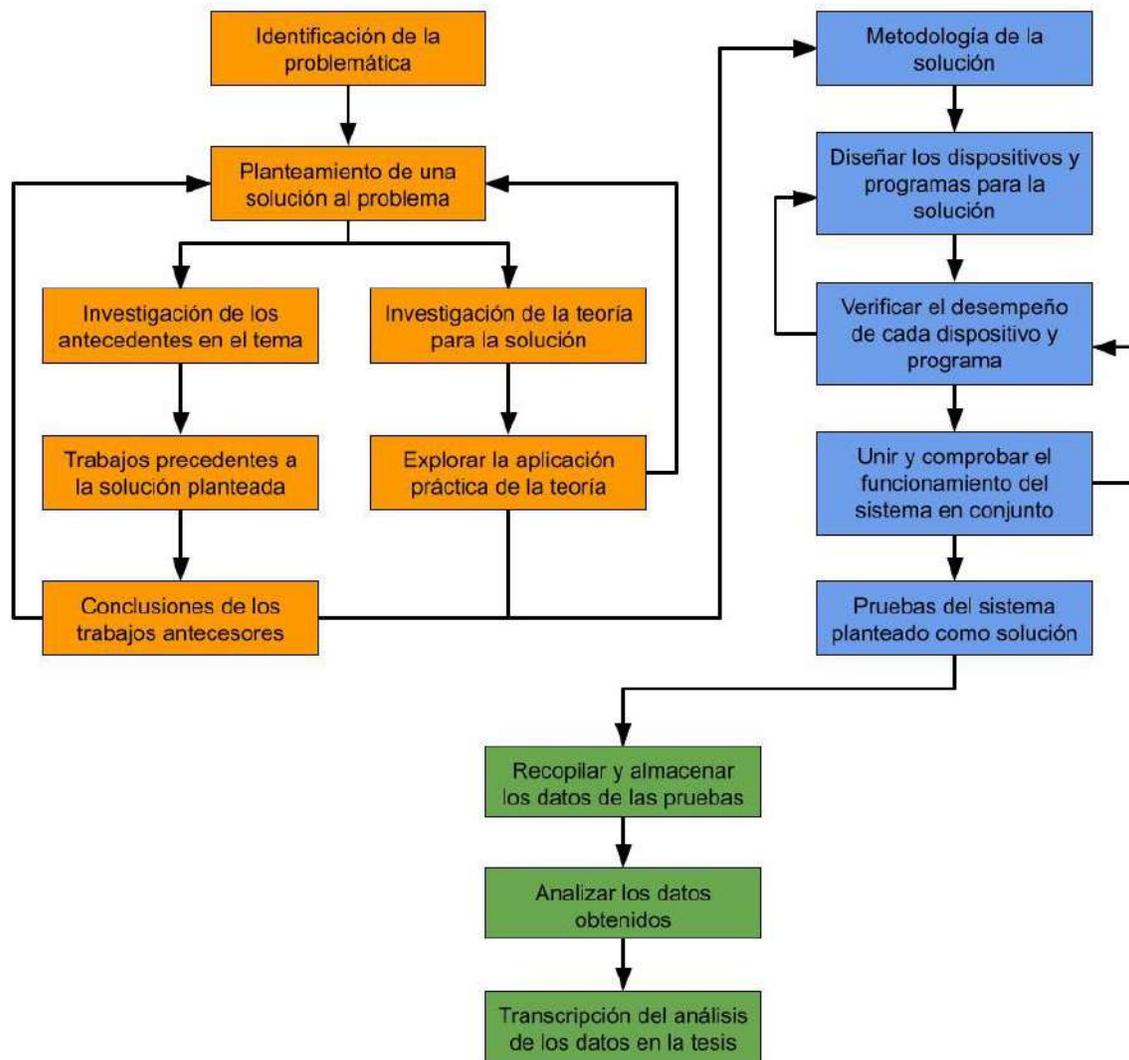


Figura 1-1. Diagrama del planteamiento general (Autoría propia, 2024).

2 FUNDAMENTACIÓN TEÓRICA

A continuación, se describen algunos fundamentos necesarios para el desarrollo del presente trabajo: el sistema de escritura y lectura braille, las redes neuronales artificiales, el significado de los OCR, el procesamiento de imágenes digitales y los diferentes métodos que se emplean, el concepto de la manufactura aditiva, los microcontroladores y las partes que lo conforman, las cámaras digitales y los actuadores a utilizar.

2.1 Sistema braille

El braille es el sistema de lectura y escritura utilizado primordialmente por las personas con discapacidad visual, fue diseñado con el propósito de ser inspeccionado de manera táctil en una celda o arreglo de puntos, en cada una de estas celdas se encuentra un grupo de seis puntos que sobresalen como un relieve y que pueden estar dispuestos en diferentes combinaciones para representar una letra del alfabeto, un número y uno de los distintos signos de puntuación (Simón et al., 1995). La Figura 2-1 muestra la forma de representar una celda braille y en la Figura 2-2 se observa el orden de posición de cada uno de los puntos que conforman la matriz braille, un carácter en este sistema se representa cuando los puntos están en negro, o en el caso de un reconocimiento por tacto estos forman un relieve, existen 64 combinaciones para una sola celda braille incluyendo cuando todos los puntos están en blanco o que la superficie no tiene relieve.



Figura 2-1. Matriz de seis puntos, distribuidos en dos columnas y tres renglones (Martínez Calvo, 2006).

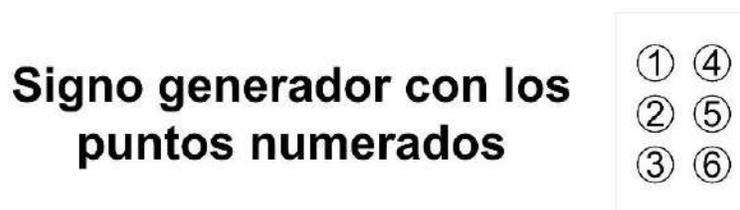


Figura 2-2. Disposición de los puntos para una celda braille (Martínez Calvo, 2006).

Las medidas que debe tener una celda braille en español se muestran en la Figura 2-3, estos tamaños los define la CBE para que una persona sea capaz de reconocer de la mejor manera un signo braille, la imagen en el inciso (a) expone las distancias que una celda tiene que cumplir para tener un tamaño adecuado; en cambio, la imagen del inciso (b) presenta las medidas para los puntos de la matriz: **a** es la distancia entre los centros de puntos contiguos de la misma celda, de 2.4 a 2.75 mm; **b** es la distancia vertical entre los centros de puntos contiguos de la misma celda, de 2.4 a 2.75 mm; **c** es la distancia entre los centros de puntos idénticos de celdas contiguas, de 6 a 6.91 mm; **d** es la distancia entre los centros de puntos idénticos de líneas contiguas, de 10 a 11.26 mm; **e** es el diámetro de la base de los puntos, entre 1.2 y 1.9 mm; **f** es la altura recomendada de los puntos, entre 0.2 y 0.5 mm, estas medidas permiten estandarizar los documentos con escritura braille (García Soria & Ruiz Prieto, 2014).

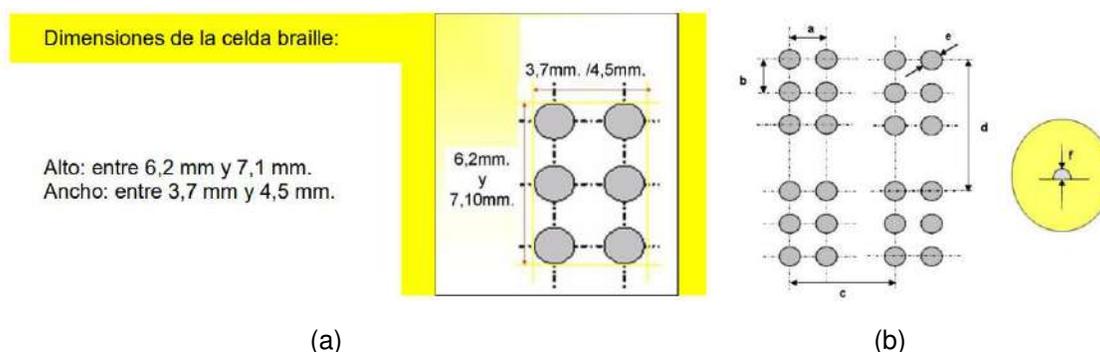


Figura 2-3. Normas de la CBE para signos braille: (a) dimensiones para una celda braille, (b) los puntos que la conforman (García Soria & Ruiz Prieto, 2014).

El idioma español tiene 27 letras en su abecedario, el braille en español contiene la misma cantidad en su alfabeto, el inciso (a) de la Figura 2-4 muestra los caracteres que lo conforman, las letras con algún signo diacrítico en español tiene su propia representación en braille, como es el caso de las vocales con acentos y de la letra *u* con diéresis, incluso la letra *ñ* entra en esta categoría, pero esta última ya está dentro del alfabeto braille español. Los números en braille utilizan los primeros 10 signos del alfabeto para representar un número del 1 al 9, y el 0, en este caso el braille hace uso de una celda especial, el signo de número, el cual indica que los siguientes signos braille posteriores a este son dígitos; el inciso (b) de la Figura 2-4 expone los ejemplos de las diez cifras. Las letras mayúsculas también usan un signo específico, este es colocado antes de un signo o palabra escrita en braille, puede señalar que la primera letra es mayúscula o que todos los caracteres de la palabra son mayúsculas; el inciso (d) muestra ejemplos de las letras mayúsculas en braille. Los signos de puntuación en español también se representan en braille, el inciso (e) de la imagen expone la forma en que estos caracteres se convierten en el sistema de escritura y lectura braille.

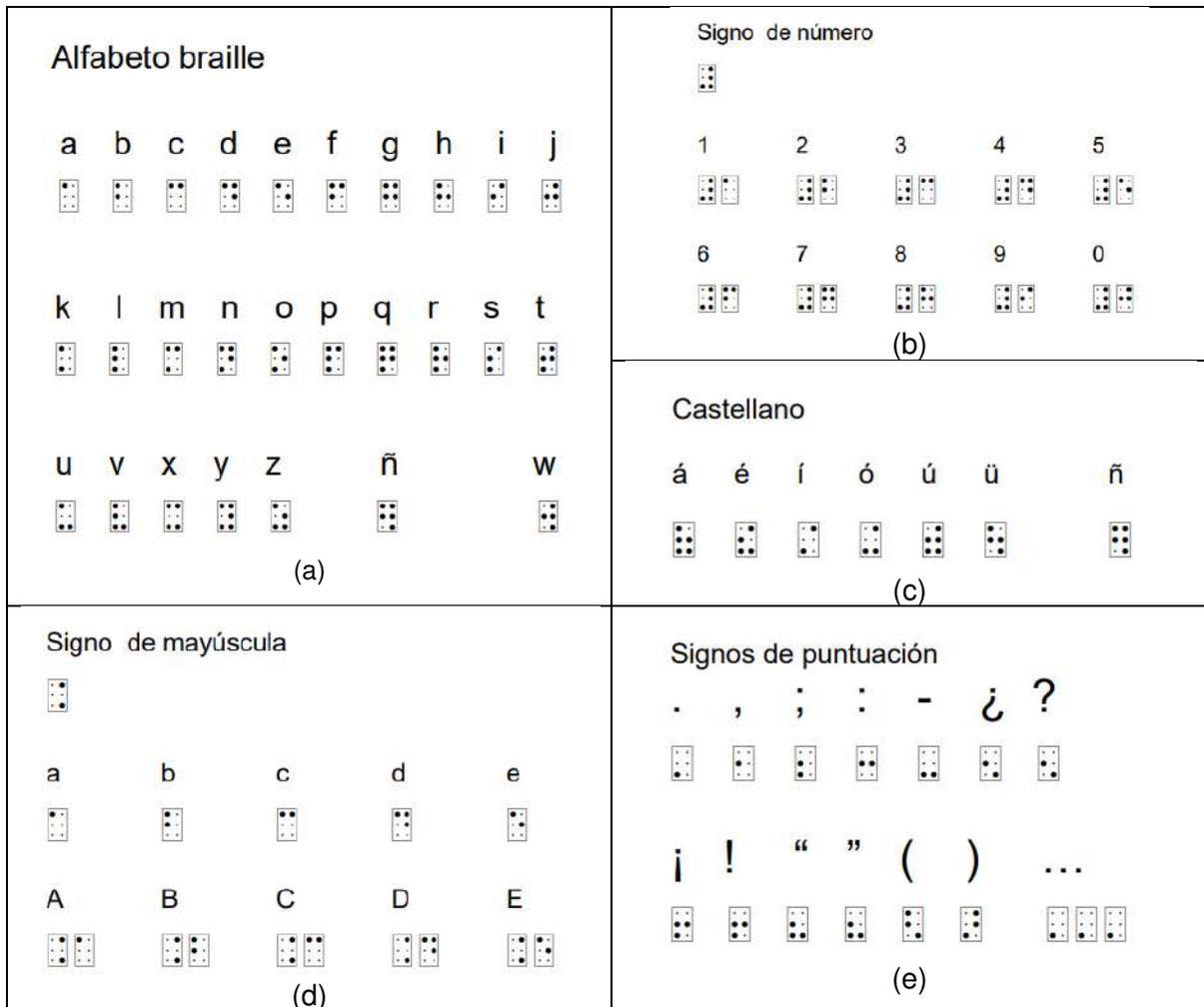


Figura 2-4. Signos del braille español: (a) alfabeto braille, (b) números en braille, (c) signos diacríticos del español en braille, (d) ejemplos de representación de mayúsculas, (e) signos de puntuación en braille (García Soria & Ruiz Prieto, 2023).

2.2 Redes neuronales artificiales

Una red neuronal artificial es una red interconectada paralelamente de neuronas artificiales simples, las cuales ejecutan procesos locales, cada una de estas recibe un conjunto de entradas de otras neuronas y obtiene un resultado que transmite por medio de un nodo de salida (Rao, 2009). Las redes neuronales artificiales tienen la facultad de obtener un resultado para problemas complejos, y que por otros métodos convencionales los cálculos serían difíciles, las redes neuronales artificiales poseen un gran potencial debido a que son capaces de aprender (Aguado Behar, 2000). Las redes neuronales artificiales se componen de neuronas simples, como la que se muestra en el inciso (a) de la Figura 2-5, este elemento también es llamado perceptrón simple por solamente tener una neurona en la red.

Los cálculos que realiza una sola neurona son llamados funciones de activación o funciones transferenciales, estas funciones están constituidas por ecuaciones algebraicas o diferenciales y pueden ser lineales o no lineales (Aguado Behar, 2000). La salida que entrega una neurona es una función, un ejemplo de esto es el inciso (b) de la Figura 2-5, generalmente la salida se genera cuando los niveles de activación sobrepasan un determinado umbral.

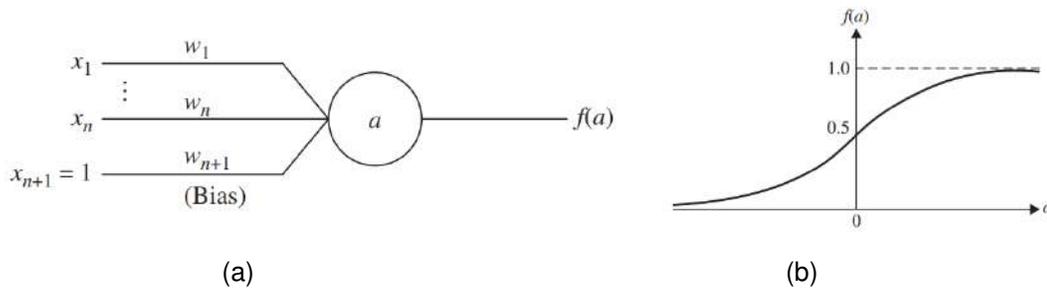


Figura 2-5. Elementos de una red neuronal artificial: (a) neurona simple, (b) una función de salida sigmoide (Rao, 2009).

El ejemplo de una red neuronal artificial se muestra en la Figura 2-6, esta red se compone de neuronas de procesamiento (círculos), líneas de conexión por donde se transfiere la información y de neuronas de almacenamiento (rectángulos) donde se guardan los valores de entrada a la red (Aguado Behar, 2000). La red que se observa en la Figura 2-6, está compuesta por distintas capas; una Capa de Entrada, que es lugar donde las neuronas guardan los datos que entran a la red; la capa oculta, la cual está compuesta por neuronas que realizan cálculos locales; y la Capa de Salida, que son las neuronas que proporcionan los resultados finales del procesamiento de toda la red.

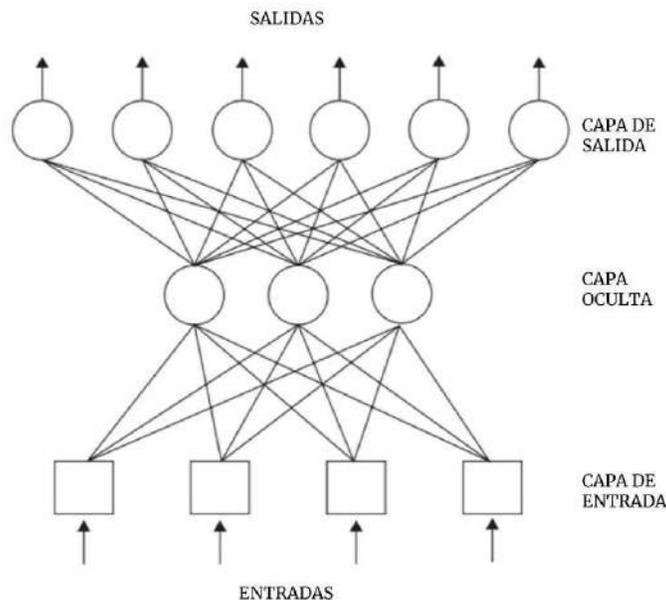


Figura 2-6. Red retroalimentada de multicapa (Rao, 2009).

Existen diferentes tipos de arquitecturas para las redes neuronales, estas dependen del número de neuronas que utiliza, las interconexiones entre las neuronas, la naturaleza del umbral en la función de salida y del proceso de aprendizaje de la red neuronal (Rao, 2009). El proceso de aprendizaje de una red se presenta mediante ejemplos, a partir de estos la red neuronal substrahe las relaciones primordiales que se encuentra implícitas en los datos guardados (Aguado Behar, 2000). Una red neuronal es entrenada para minimizar el error cuadrático entre los valores que entregan las salidas y los datos almacenados, el error disminuye cuando los pesos son ajustados en varias líneas de conexión (Rao, 2009).

2.3 Procesamiento de imágenes digitales

El procesamiento de una imagen por medio de una computadora es la forma de realizar este tipo denominado proceso, para realizar un determinado tipo de operaciones computacionales en una imagen ésta es definida como una función bidimensional, $f(x, y)$, en donde x y y son coordenadas en el plano, y la amplitud de f para cualquier par de coordenadas (x, y) es nombrada la intensidad o el nivel de gris en punto preciso de la imagen. La imagen digital está compuesta por un número finito de componentes, los cuales tienen valores finitos para x , y y el nivel de gris, estas partes que componen la imagen son denominados elementos pictóricos o píxeles (Gonzalez & Woods, 2018). Los seres humanos tienen la visión como el sentido corporal que permite percibir el mundo a través de la luz, el espectro electromagnético que somos capaces de observar es solamente una fracción, pero los diferentes tipos de cámaras y dispositivos electrónicos dedicados a percibir la luz cubren un rango mucho más grande en el espectro electromagnético.

El procesamiento de imágenes digitales está conformado por dos categorías principales, en una las entradas y salidas son imágenes, mientras que en la otra se mantiene como entradas las imágenes, pero las salidas que entregan son atributos extraídos de las propias imágenes, esta forma de organización de procesos es mostrada en el diagrama de la Figura 2-7. El diagrama presenta un compendio de las diferentes metodologías que pueden ser aplicadas en imágenes con diferentes propósitos (Gonzalez & Woods, 2018).

La función que realizan o debe ejecutar cada uno de los procesos mostrados en la Figura 2-7 son los siguientes:

- **Adquisición de imagen.** Es el primer proceso donde una imagen simple es recibida de forma digital, en este paso generalmente está involucrado el proceso de escalado.



Figura 2-7. Pasos fundamentales en el procesamiento de imágenes digitales (Gonzalez & Woods, 2018).

- **Filtrado y mejora de imágenes.** Es el procesamiento de una imagen en donde tal manipulación entrega un resultado con mejoras en la imagen para una aplicación en concreto. El mejoramiento de una imagen no tiene una teoría en general, se basa más en el juicio de un espectador y que él mismo determine función del método aplicado, además existen una gran variedad de técnicas para mejorar la fotografía.
- **Restauración de imagen.** Esta área también se dedica a la mejora de una imagen, el proceso de restauración de imágenes es objetivo, porque esta técnica tiende a fundamentarse en modelos matemáticos y probabilísticos de degradación, a diferencia de la mejora de imagen, la cual es subjetiva.
- **Procesamiento de imágenes de color.** En esta metodología el color es la base para la extracción de características de una imagen, y los fundamentos de esta técnica son los modelos de color y cómo estos son procesados en el dominio digital.
- **Ondículas y otras transformaciones de imágenes.** Es la base para representar imágenes con sus diferentes grados de resolución, así como en el tema de información comprimida en imágenes y de representaciones piramidales, en las cuales se divide una imagen en una zona más pequeña.

- **Compresión y marca de agua.** El proceso está vinculado con técnicas que reducen el almacenamiento requerido para guardar una imagen, o del ancho de banda necesario para que esta sea transmitida.
- **Procesamiento morfológico.** Esta metodología trabaja con herramientas que extraen elementos de una imagen, los cuales son útiles para reproducir y describir una forma.
- **Segmentación.** Es la división de una imagen en partes que la conforman, este procesamiento requiere de un largo camino para obtener buenos resultados, ya que esta metodología se encarga de solucionar problemas en la imagen de manera particular.
- **Extracción de características.** Este método consiste en el reconocimiento y descripción de características en una región o límite de la imagen, esta técnica designa valores cuantitativos a las características detectadas. Esta metodología se relaciona con el proceso de segmentación, porque frecuentemente la salida de este último es la entrada de la técnica de extracción.
- **Clasificación de patrones en imágenes.** Es el proceso que determina un título para una imagen dependiendo de las descripciones y características obtenidas.

❖ Representación de imágenes digitales

Las imágenes adquiridas en un formato digital son representadas de manera general en matrices como se muestra en la Figura 2-8, el inciso (a) es la forma en que se presentan las imágenes en un monitor de computadora o una fotografía, en cambio el inciso (b) muestra la imagen digital como una matriz compuesta de valores numéricos (Gonzalez & Woods, 2018).

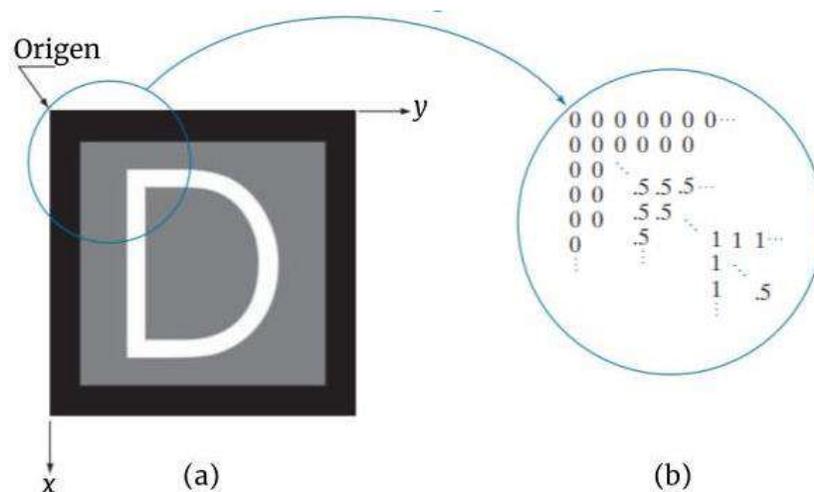


Figura 2-8. Representación de imágenes digitales: (a) imagen presentada como una matriz de intensidad visual, (b) imagen mostrada como una matriz numérica (Gonzalez & Woods, 2018).

La fotografía que es contenida en $f(s, t)$ representa una función de imagen continua, con variables s y t también continuas. La conversión al formato digital de una imagen se realiza mediante el muestreo y cuantificación, de esta manera la función cambia a ser $f(x, y)$, donde esta misma tiene un número M filas y un número N de columnas, y las variables x y y son coordenadas. La representación de una imagen digital en forma de matriz es igual a la ecuación (1), los valores para cada una de las variables son enteros: $x = 0, 1, 2, \dots, M$, $y = 0, 1, 2, \dots, N$ y los cuales simbolizan coordenadas en la imagen digital; de este modo, el valor en el origen es denotado como $f(0,0)$, y para el siguiente valor en la primera fila es $f(0,1)$. Con esto, el valor para una imagen digital es denotado como $f(x, y)$ en cualquiera de las coordenadas de (x, y) (Gonzalez & Woods, 2018).

$$f(x, y) = \begin{pmatrix} f(0,0) & f(0,1) & \dots & f(0, N - 1) \\ f(1,0) & f(1,1) & \dots & f(1, N - 1) \\ \vdots & \vdots & \dots & \vdots \\ f(M - 1, 0) & f(M - 1, 1) & \dots & f(M - 1, N - 1) \end{pmatrix} \quad (1)$$

La representación de una imagen en forma de ecuación permite realizar el procesamiento por computadora a la misma, los valores que la función f almacena en cada una de sus coordenadas pueden ser entre una escala de grises o para guardar los datos de una imagen a color. En una escala de grises, el valor de intensidad que $f(x, y)$ guarda es entre 0 y 1, simbolizando 0 el color negro, el 1 el color blanco y una intensidad de gris puede estar entre ambos valores; un ejemplo de esto se observa en la Figura 2-8: el inciso (a) presenta una imagen en negro, gris y blanco; el inciso (b) muestra los valores 0, 0.5 y 1 para los colores negro, gris y blanco respectivamente en una matriz, tal y como estarían ordenados en la imagen del primer inciso. Para una imagen a color, los valores que se almacenan en cada una de sus coordenadas también son de diferentes intensidades de color, pero en este caso son para tres colores: rojo, verde y azul; en la ecuación (2) se puede observar que $f(x, y)$ guarda los datos en forma de un vector, donde $f_R(x, y)$ es un valor del color rojo (R, sigla de “rojo” en inglés), $f_G(x, y)$ del verde (G, sigla de “verde” en inglés) y $f_B(x, y)$ del azul (B, sigla de “azul” en inglés), estos mismo valores también pueden estar entre una cantidad de 0 a 1; por ejemplo, en, en una coordenada (x, y) el valor de $f(x, y) = [1, 0, 0]$ representa el color rojo, $f(x, y) = [0, 1, 0]$ el verde y $f(x, y) = [0, 0, 1]$ el azul; diferentes cantidades en cada una de las componentes de $f(x, y)$ representa un determinado color o una combinación que estos generan (Gonzalez & Woods, 2018).

$$f(x, y) = \begin{bmatrix} f_R(x, y) \\ f_G(x, y) \\ f_B(x, y) \end{bmatrix} = \begin{bmatrix} R(x, y) \\ G(x, y) \\ B(x, y) \end{bmatrix} \quad (2)$$

❖ Ejemplos de procesamiento de imágenes

En el procesamiento de imágenes, la fotografía que se adquiere por medio de una cámara digital es generalmente almacenada como una imagen a color, la información es guardada de la misma forma que la ecuación (2), teniendo en cada una de las coordenadas de la imagen un valor para el color rojo (R), verde (G) y azul (B), con estos datos es posible generar diferentes procesamientos, entre los más comunes esta la conversión de la imagen a una escala de grises y la otra es a blanco y negro.

• Conversión de una imagen de color a una escala de grises

La transformación a una escala de grises de una imagen a color se realiza procesando los valores de cada coordenada de la imagen en la ecuación (3), en este caso la información esta almacenada en la forma de la ecuación (1) y en cada coordenada (x, y) los colores están guardados igual que la ecuación (2).

$$g(x, y) = 0.3R(x, y) + 0.59B(x, y) + 0.11G(x, y) \quad (3)$$

En el momento que la ecuación (3) realiza operaciones con los valores de los colores en una coordenada de la imagen, esta genera un dato entre 0 y 1, este mismo es un valor para la intensidad de gris de tal coordenada. El proceso se repite $M \times N$ veces hasta que cada una de las coordenadas de la imagen es transformada, en consecuencia, se genera la ecuación (4), en la que los datos generados se encuentran entre 0 y 1, o en este caso con diferentes intensidades del color gris.

$$g(x, y) = \begin{pmatrix} g(0,0) & g(0,1) & \dots & g(0, N - 1) \\ g(1,0) & g(1,1) & \dots & g(1, N - 1) \\ \vdots & \vdots & \dots & \vdots \\ g(M - 1,0) & g(M - 1,1) & \dots & g(M - 1, N - 1) \end{pmatrix} \quad (4)$$

• Conversión de una imagen de color a blanco y negro

La conversión a blanco y negro de una imagen a color requiere que primero se haya realizado la conversión de colores a escalas de grises, con esto obtenido, la transformación a blanco y negro solo requiere de la ecuación (5).

$$h(x, y) = \begin{cases} 1 & \text{umbral} \geq g(x, y) \\ 0 & \text{umbral} < g(x, y) \end{cases} \quad (5)$$

El procesamiento que emplea la ecuación (5), es simplemente comparar la información que se encuentra en cada una de las coordenadas de la ecuación (4) con un valor de intensidad de gris, en este caso un “umbral”, si la cantidad en una determinada coordenada (x, y) es mayor o igual al umbral, el dato que se genera es 1, pero si es menor, el dato es 0; esta comparación se repite en todas las coordenadas hasta generar una matriz similar a la ecuación (6), y en donde los valores que ésta matriz tiene son 0 y 1 o el color negro y blanco respectivamente.

$$h(x, y) = \begin{pmatrix} h(0,0) & h(0,1) & \dots & h(0, N - 1) \\ h(1,0) & h(1,1) & \dots & h(1, N - 1) \\ \vdots & \vdots & \dots & \vdots \\ h(M - 1,0) & h(M - 1,1) & \dots & h(M - 1, N - 1) \end{pmatrix} \quad (6)$$

2.4 Reconocimiento Óptico de Caracteres

El Reconocimiento Óptico de Caracteres (OCR por sus siglas en inglés) es el proceso de transformación de documentos de texto impresos o escritos a mano, en símbolos legibles para una computadora, como archivo de texto, HTML, entre otros (Borovikov, 2004). El sistema de reconocimiento OCR consiste en tres grandes etapas:

1. Digitalización del documento
2. Reconocimiento de caracteres o palabras
3. Distribución de salida

En la primera etapa los documentos con texto impreso o escritos a mano pasan por un escaneo para generar imágenes de los propios documentos, la Figura 2-9 muestra el ejemplo con un documento que es colocado en un escáner. La segunda etapa registra los caracteres o palabras que están contenidos en las imágenes de los documentos, este proceso involucra cuatro operaciones:

- Análisis de la imagen: evaluar la calidad de la captura, detección de líneas de texto, extracción de caracteres y palabras.
- Mejora de la imagen: remover ruidos en la imagen, rellenar huecos y rupturas, entre otras mejoras.
- Reconocimiento de los caracteres y palabras basándose en sus formas y otras características.
- El procesamiento contextual para limitar el espacio de búsqueda de las características.

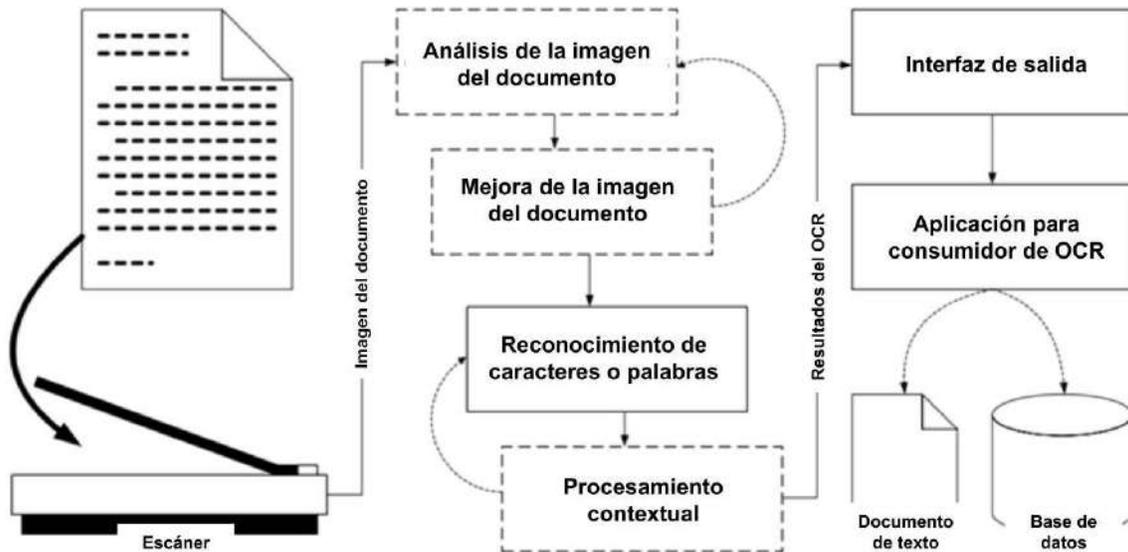


Figura 2-9. Sistema OCR típico (Borovikov, 2004).

La tercera y última etapa del sistema OCR entrega los resultados por medio de una interfaz de salida para ser expuesto en una aplicación, dependiendo de la función de esta última la información es almacenada en archivos para hojas de cálculo, para procesadores de texto e incluso en bases datos.

Los sistemas OCR contienen dos componentes esenciales: el extractor de características y un clasificador. El extractor es utilizado como una entrada al clasificador y determina los elementos conocidos que corresponden con las características observadas. El clasificador entrega un número de nivel de confianza en cuanto a la seguridad de estar reconociendo el elemento (Borovikov, 2004).

2.5 Manufactura Aditiva

La Manufactura Aditiva (AM, *Additive Manufacturing*) es una manera de referirse a la recreación rápida de prototipos, y es nombrado comúnmente como impresión en 3D; el término de recreación rápida de prototipos es utilizado por fabricantes industriales como un proceso que recrea rápidamente partes de un sistema. La manufactura aditiva tiene énfasis en crear en el menor tiempo posible prototipos, los cuales sirven de modelos de pruebas para lo que posteriormente son los productos comerciales (Gibson et al., 2021). Un ejemplo de este proceso de manufactura se muestra en la Figura 2-10, en ella se puede observar el esquema de un sistema de manufactura aditiva por extrusión; el funcionamiento de este consiste en que sobre una plataforma se coloca material de filamento, el

cual primero es arrastrado a presión a una cámara licuadora, en esta misma el material se calienta hasta llegar a un estado líquido, y una vez que el material es líquido sale a través de la punta de una boquilla para ser depositado en la plataforma; tanto la cámara licuadora y la punta se mueven en los ejes x-y y z para colocar el material en forma de andamio, y de esta forma manufacturar una pieza en tres dimensiones.

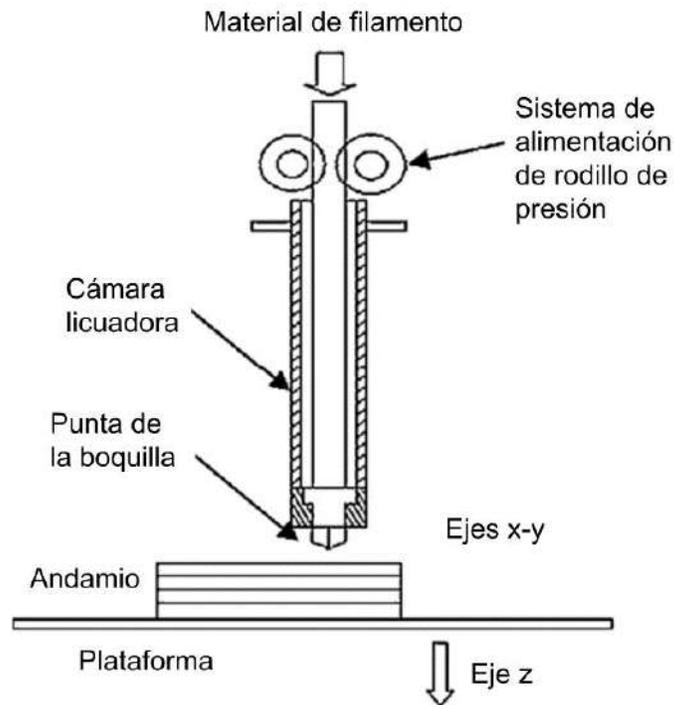


Figura 2-10. Esquema de un sistema basado en extrusión (Gibson et al., 2021).

En términos más estandarizados la ASTM (Sociedad Estadounidense para Pruebas y Materiales, *American Society for Testing and Materials*) y la ISO (Organización Internacional de Normalización, *International Organization for Standardization*) definen la manufactura aditiva de la siguiente manera: “*La manufactura aditiva es el proceso de unión de materiales para crear partes a partir de la información de modelos 3D, comúnmente capa sobre capa, y en contraste con las metodologías de manufactura por sustracción y de fabricación formativa*” (Toyserkani et al., 2022).

La funcionalidad de una pieza manufacturada depende de las propiedades y geometría del objeto a fabricar, la combinación de estos dos aspectos permite que los procesos de fabricación den forma a una pieza de un material en concreto, mediante operaciones y subprocesos. La creación de objetos con el uso de materiales en un proceso de manufactura se logra mediante la implementación de uno, o de la combinación de tres principios básicos: formación formativa, formación sustractiva y formación aditiva (Toyserkani et al., 2022). La manufactura aditiva esta categorizada de manera estándar en los siguientes siete procesos:

1. Chorro aglomerante (*Binder Jetting*)
2. Deposición de energía directa (*Direct Energy Deposition*)
3. Extrusión de materiales (*Material Extrusion*)
4. Chorro de material (*Material Jetting*)
5. Fusión de lecho de polvo (*Powder Bed Fusion*)
6. Laminación de láminas (*Sheet Lamination*)
7. Fotopolimerización (*Photopolymerization*)

2.6 Microcontroladores

Un microcontrolador es un microcomputador que se forma en un circuito integrado (chip) para ser implementados en una gran variedad de aplicaciones y con propósitos distintos: en automoción, equipos de comunicaciones y telefonía, instrumentos electrónicos, equipos médicos, equipos industriales, en electrodomésticos, juguetes, entre otro tipo de aparatos. Los microcomputadores están conformados por una unidad central de procesamiento (CPU) o un microprocesador, una memoria y los elementos de entrada y salida, estas tres partes están conectadas por medio de líneas llamadas buses, la Figura 2-11 muestra el diagrama general de un microcomputador (Valdéz Pérez & Pallás Areny, 2007).

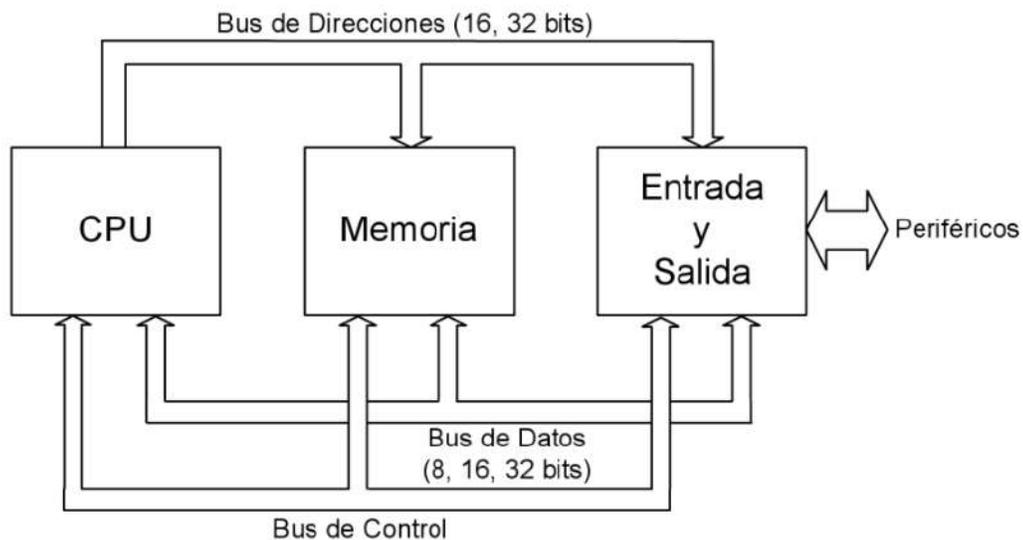


Figura 2-11. Esquema general de un microcomputador (Valdéz Pérez & Pallás Areny, 2007).

Los microcontroladores también están formados por componentes que posee un microcomputador; tienen una CPU, una memoria y las entradas y salidas de un circuito integrado, aunque contienen otros elementos tal y como se muestra

en la Figura 2-12. Entre los elementos principales del microcontrolador está el oscilador, el cual se encarga de sincronizar todos los cálculos que tiene el circuito interno, un oscilador que determina la frecuencia de la velocidad de ejecución de las ordenes en el microcontrolador (Valdéz Pérez & Pallás Areny, 2007).

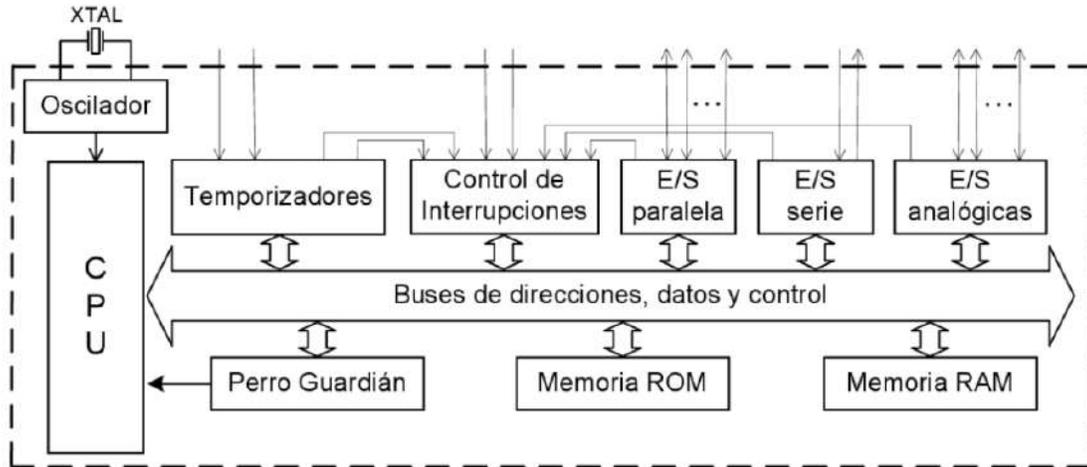


Figura 2-12. Esquema de bloques general de un microcontrolador (Valdéz Pérez & Pallás Areny, 2007).

El CPU es el componente que ejecuta las órdenes del microcontrolador, estas son extraídas individualmente desde la memoria, las descifra e implementa. El CPU contiene varios registros para propósitos específicos: un Registro de Instrucción (RI) que almacena las órdenes que se están ejecutando; el Acumulador (ACC, *Accumulator*) que está ligado a operaciones aritméticas y lógicas; el Registro de Estado (STATUS) es el encargado de agrupar los bits indicadores de resultados en las operaciones aritméticas y lógicas; el Contador de Programa (PC, *Program Counter*) es el registro contenedor de las direcciones para las órdenes a ejecutar; el Registro de Direcciones de Datos (RDD) guarda en la memoria las direcciones de datos; y el Puntero de la Pila (SP, *Stack Pointer*), el cual toma las direcciones de datos de pila y las almacena (Valdéz Pérez & Pallás Areny, 2007).

El microcontrolador alberga dos tipos de memoria: la memoria RAM (*Random Access Memory*) y la memoria ROM (*Read Only Memory*). La memoria ROM se utiliza para almacenar programas a ejecutar de manera permanente, la ROM realiza solamente el trabajo de lectura, mientras que la RAM es una memoria tanto de lectura como escritura, pero esta última solo resguarda de manera temporal la información del programa con el que está trabajando. La memoria RAM es volátil, pierde los datos resguardados al momento de no ser alimentada con energía, en cambio, la memoria ROM no es volátil y es por esta razón es que mantiene los datos permanentemente incluso cuando no se le suministra energía (Valdéz Pérez & Pallás Areny, 2007).

Las entrada y salidas son elementos de un microcontrolador que le permiten comunicarse con el exterior, estos componentes están conformados por puertos paralelos, puertos en serie, los temporizadores, la gestión de interrupciones, además pueden incluirse las entradas y salidas analógicas vinculadas a un Convertidor Analógico Digital (ADC, *Analog-to-Digital Converter*) y Digital Analógico (DAC, *Digital-to-Analog Converter*). Los puertos en serie pueden ser de diferente tipo dependiendo del protocolo de comunicación que utilicen: RS-232C, USB, Ethernet, entre otros. Los elementos de entrada y salida utilizan un mismo conjunto de terminales, una misma terminal es conectada a distintos bloques de entrada y salida de manera interna (Valdéz Pérez & Pallás Areny, 2007).

El perro guardián (WDT, *Watchdog Timer*) es un componente encargado de reiniciar el microcontrolador después de un intervalo de tiempo específico, este elemento hace un conteo de los pulsos provenientes del oscilador, y llegado a cierto número de pulsos el WDT activa una salida que reestablece el microcontrolador (Valdéz Pérez & Pallás Areny, 2007).

2.7 Cámaras digitales

Las cámaras digitales son dispositivos que adquieren la luz, transforman los fotones en electrones; con el pasar del tiempo las cámaras digitales mismas se han vuelto más pequeñas, igualmente se ha mejorado la forma de capturar los fotones y almacenarlos en electrones, una de estas mejoras es mediante el uso de silicón. El silicón es un semiconductor sensitivo a la luz, en la ausencia de luz el silicón es una aislante mientras que con su presencia éste se convierte en conductor, y de esta manera es capaz de transformar la luz en electrones para ser almacenados (Griffing, 2023).

En las cámaras digitales el diseño para los sensores es similar a una disposición cuadrada de los pixeles en una imagen digital. La distribución de los elementos del sensor permite almacenar los fotoelectrones cuando está expuesto a la luz durante un periodo de tiempo, además la forma en que se guarda la información para representar exactamente la imagen depende del sensor utilizado; un tipo de cámaras implementan un sensor CCD (Dispositivo de Carga Acoplada, *Charge-Coupled Device*), mientras que otras utilizan el CMOS (Semiconductor de Óxido de Metal Complementario, *Complementary Metal Oxide Semiconductor*), la Figura 2-13 muestra la forma del arreglo de los dos sensores CCD y CMOS (Griffing, 2023).

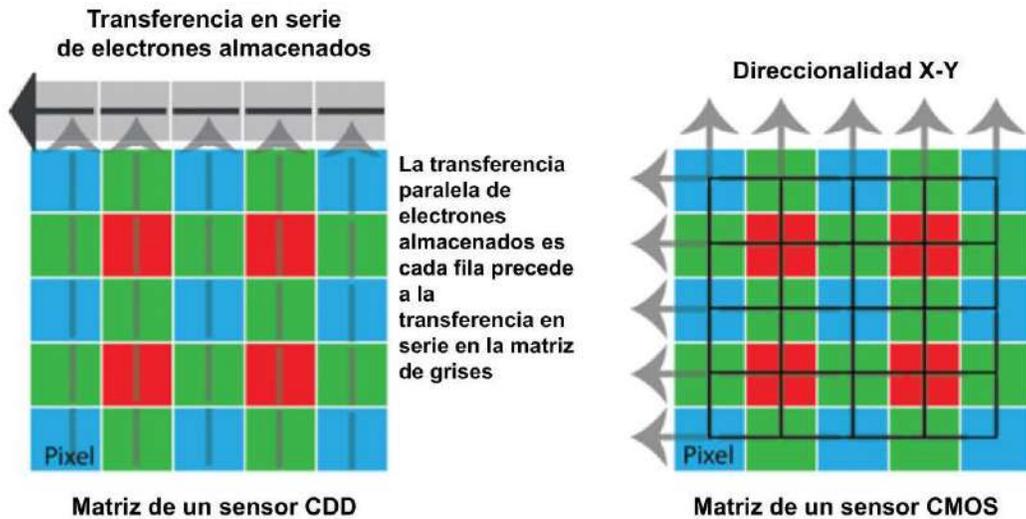


Figura 2-13. Diseño de los sensores CDD y CMOS para una cámara digital (Griffing, 2023).

Los sensores CMOS son implementados por cámaras de uso comercial, esto se debe a su velocidad para capturar imágenes de video, teniendo un rango de 20 a 30 fotogramas por segundo (FPS, *Frames Per Second*), incluso llegando a más de 100 fotogramas por segundo dependiendo del sensor; aunque los sensores CMOS producen el efecto de ruido de patrón fijo, este es un problema del muestreo y la conmutación para identificar cada píxel, pero los nuevos sensores CMOS han superado este problema. En los sensores CDD existe una mayor facilidad de fabricación, la lectura de cada píxel tiene menos ruido y requieren un menor procesamiento, pero el rango de fotogramas por segundo para el CCD se encuentra entre 5 a 10 fotogramas por segundo (Griffing, 2023).

2.8 Actuadores solenoides

Los solenoides como un sistema de actuación son prácticamente un circuito magnético, en donde una bobina es colocada alrededor de un émbolo de material ferromagnético, cuando por este circuito se induce una corriente eléctrica, el campo que genera la bobina provoca que el émbolo se desplace a una posición de menor reluctancia, el movimiento del émbolo es provocado por el flujo magnético y se coloca en un lugar donde ofrezca una menor resistencia a este mismo flujo (Gomis-Bellmunt & Campanile, 2010).

La forma de un solenoide se muestra en la Figura 2-14, la imagen del inciso (a) expone un ejemplo de lo que es un bosquejo para fabricar un solenoide, en éste se observa que el solenoide tiene un émbolo rodeado por cables de cobre y, estos mismos se encuentran cubiertos por dos láminas y un tubo de hierro; mientras en la imagen en el inciso (b) se muestran varios ejemplos de solenoides industriales.

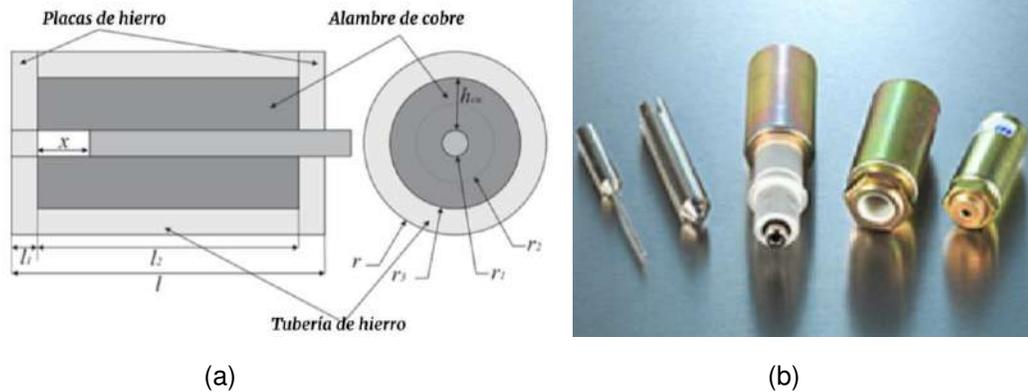


Figura 2-14. Imágenes de solenoides: (a) bosquejo de un actuador solenoide, (b) solenoides industriales (Gomis-Bellmunt & Campanile, 2010).

2.9 Software e interfaz

El IEEE (Instituto de Ingenieros Eléctricos y Electrónicos, *Institute of Electrical and Electronics Engineers*) define la palabra “software” de la siguiente manera: *Es el conjunto de los programas de cómputo, procedimientos, reglas, documentación y datos asociados, que forman parte de las operaciones de un sistema de computación* (IEEE, 1990).

El significado de “software” que proporciona el IEEE puede que no siempre este acorde a lo que un profesional de ingeniería en software entiende, incluso la mayoría los usuarios pueden tener una idea distinta. Las definiciones de “software” que ofrece un libro serian de la siguiente manera: 1) son instrucciones (programas de cómputo) que cuando se ejecutan proporcionan las características, la función y el desempeño buscado; 2) son estructuras de datos que permiten que los programas manipulen en forma adecuada la información; 3) es la información descriptiva tanto en papel como en forma virtual, la cual describe las operaciones y usos de los programas (Pressman, 2010). El software es parte de un sistema lógico y tiene diferentes características a las del hardware o un sistema físico:

1. El software es desarrollado de manera intelectual y su modificación requiere que la persona encargada del desarrollo conozca bien el diseño del producto construido.
2. El software no se puede “dañar” en un sentido físico, esto se relaciona más con el hardware, el cual puede llegar a desgastarse, pero hacer una modificación en el software o realizar cambios en él puede generar fallos que lo conducen a un deterioro; además la constatación de modificación para solucionar errores propicia que el mantenimiento del software sea complejo.

3. En la industria del software los desarrollos son creados para ser utilizados de manera personal, pero las partes que conforman una nueva aplicación en software pueden ser extraídas de librerías con componentes reusables.

❖ **La interfaz de usuario**

La interfaz es un medio por el cual los seres humanos pueden comunicarse con las computadoras de una forma eficaz, esta misma debe seguir fundamentos de diseño para tener una buena “usabilidad”: esto último es una medición cualitativa de la forma que una persona trabaja fácil y eficientemente con las funciones y características de una aplicación de software. Durante una parte en la historia de la computación, la usabilidad no fue prevista como un problema a tratar por quienes fabricaban software (Pressman, 2010).

En el transcurso del tiempo se estudió la manera en que el humano interactúa con la tecnología, de esto mismo surgieron dos aspectos en los desarrollos tecnológicos. El primero fue definir un conjunto de reglas, las cuales deben ser aplicadas a productos enfocados en la interacción de las personas con la tecnología. En el segundo aspecto se determinaron un grupo de mecanismos de interacción, estos permitieron a los diseñadores crear sistemas que utilizan las reglas de una forma adecuada; los mecanismos fueron nombrados colectivamente como: Interfaz Gráfica del Usuario (GUI, *Graphical User Interface*) (Pressman, 2010). Las reglas definidas para los desarrollos tecnológicos son las siguientes:

1. **Dejar el control al usuario.**

- Definir modos de interacción de manera que no se obligue al usuario a realizar acciones innecesarias o no deseadas.
- Dar una interacción flexible.
- Permitir que la interacción del usuario sea interrumpible y también reversible.
- Ocultar los tecnicismos internos al usuario ocasional.

2. **Reducir la carga de memoria del usuario.**

- Reducir la demanda de memoria de corto plazo.
- Hacer que lo preestablecido sea significativo.
- Definir atajos que sean intuitivos.
- Revelar información de manera progresiva.

3. **Hacer que la interfaz sea consistente.**

- Permitir que el usuario coloque la tarea en curso en un contexto significativo.
- Mantener la consistencia en toda la familia de aplicaciones.
- Si los modelos interactivos anteriores han creado expectativas en el usuario, no haga cambios a menos de que haya una razón ineludible para ello.

3 METODOLOGÍA

La construcción del sistema de traducción se explica a continuación, se expone el diseño general del sistema y la metodología que este mismo sigue para su funcionamiento. También se explican las partes del sistema, los componentes que contienen, la forma en que están construidas y la manera en cómo funcionan.

3.1 Diseño general del sistema

El planteamiento general de cómo funciona el sistema de traducción de texto impreso en español a braille es mostrado en la Figura 3-1. Primeramente, se establece que la adquisición de fotos con texto impreso se realice por medio de una minicámara colocada en un soporte, en este mismo se coloca la punta de un dedo índice; el soporte de la minicámara junto con el dedo se posiciona sobre una superficie plana con texto impreso, después este realiza un movimiento horizontal para detectar diferentes caracteres de un renglón, esto último porque la minicámara debe estar continuamente tomando fotos. Las imágenes capturadas serán enviadas a una computadora, en esta misma una aplicación procesará las fotografías por medio de métodos de procesamiento de imágenes, un motor OCR y una red neuronal artificial; esto para identificar caracteres de texto español, almacenarlos de manera digital acorde al código ASCII, y después convertirlos en valores que se puedan representar en signos braille, para al final transferirlos a una tarjeta con un microcontrolador. Por último, los datos enviados a la tarjeta son interpretados por la misma, para que esta controle la activación de los solenoides, los cuales se moverán o permanecerán inmóviles acorde al carácter a representar en una matriz; esta matriz de solenoides es construida acorde a la distribución definida por la CBE para dar forma a los signos braille.

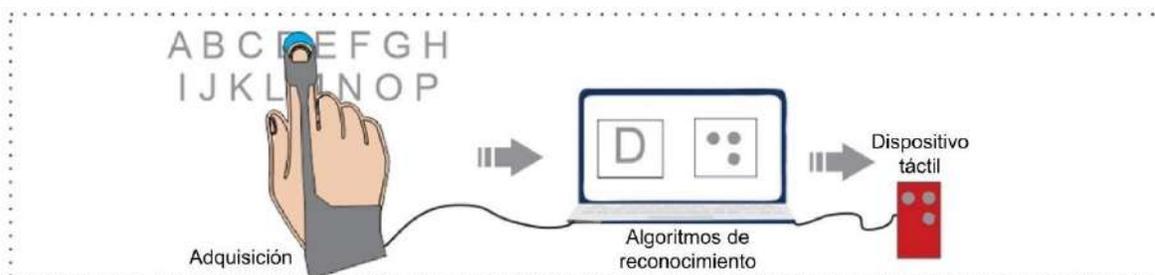


Figura 3-1. Diagrama general del sistema de traducción (Autoría propia, 2024).

La minicámara debe estar sobre un soporte en el cual se coloca la punta de un dedo índice del usuario, este soporte será sencillo en principio para no incomodar. El funcionamiento del dispositivo que representa los signos braille se realizará visualmente por el propio desarrollador del sistema, esto para comprobar que los caracteres de texto son transformados y representados en braille.

3.2 Metodología del sistema

La metodología que sigue el sistema de traducción se muestra en la Figura 3-2 de manera simple y en la Figura 3-3 de manera más extensa. En la Figura 3-2 se observa la metodología simple dividida en tres bloques: el primero consiste en el dispositivo de adquisición, que es la minicámara digital y su controlador; el segundo bloque tiene una computadora, en ella se encuentra la etapa de procesamiento de las imágenes para obtener las características y elementos de estas mismas, después almacenar el texto en la imagen digitalmente y al final convertirlo a valores que correspondan a los signos braille de cada letra; en el último bloque se encuentra el microcontrolador y el mecanismo de la celda braille donde se representan los signos brilles del texto adquirido en las imágenes.



Figura 3-2. Diagrama metodológico simple (Autoría propia, 2024).

En el diagrama metodológico de la Figura 3-3 se describe de manera más detallada el proceso por el cual se transformarán los caracteres impresos a signos braille; éste también se encuentra dividido en tres bloques generales: la etapa de adquisición, de procesamiento y de salida. La captura de la imagen, el tratamiento de esta misma y la transformación de los datos se describe en los siguientes pasos de la metodología:

- Primero el soporte para la minicámara debe estar sobre una superficie, con esto la minicámara es encendida en conjunto con leds que asisten en la iluminación.
- El siguiente paso consiste en que la minicámara tome una foto de la superficie sobre la que está colocada.
- La fotografía pasa por una tarjeta con un microcontrolador para ser enviada a una computadora.
- La imagen en la computadora es almacenada y procesada para ser convertida a blanco y negro.
- Después se realiza un procesamiento para identificar figuras que tengan relación con las formas de las letras y palabras.
- En el caso de no detectar las figuras el sistema volverá a tomar otra fotografía, pero si existen tales formas, entonces la imagen pasa al siguiente proceso.
- El siguiente paso consiste en numerar las figuras detectadas y el orden horizontal de las mismas, ya que se requiere saber cuántos caracteres en la

imagen, pueden leerse de forma horizontal, esto mismo se debe a que el movimiento del dedo índice con la minicámara imita la forma en que una persona, con alguna incapacidad para ver, lee un renglón.

- Después se ejecuta el reconocimiento de caracteres con la ayuda de un motor OCR y de esta forma se extraen las letras en la imagen.
- En caso de que no se identifique un carácter en la imagen, se envía una orden de tomar otra fotografía y repetir los pasos anteriores con la nueva información adquirida. Si se detectaron caracteres o palabras en la imagen, estos son almacenados.
- El siguiente paso consiste en convertir los caracteres o palabras al sistema braille de acuerdo con las reglas del idioma español.
- El texto en braille es codificado y enviado al microcontrolador, éste descifra los datos transferidos y los transforma en señales eléctricas.
- Las señales de salida del microcontrolador entran en un circuito de potencia para accionar los solenoides, estos se moverán de una posición a otra dependiendo de la energización de cada solenoide.
- Los solenoides están dispuestos en una matriz de tres renglones y dos columnas, la activación y desactivación de los solenoides permite representar un solo carácter braille.

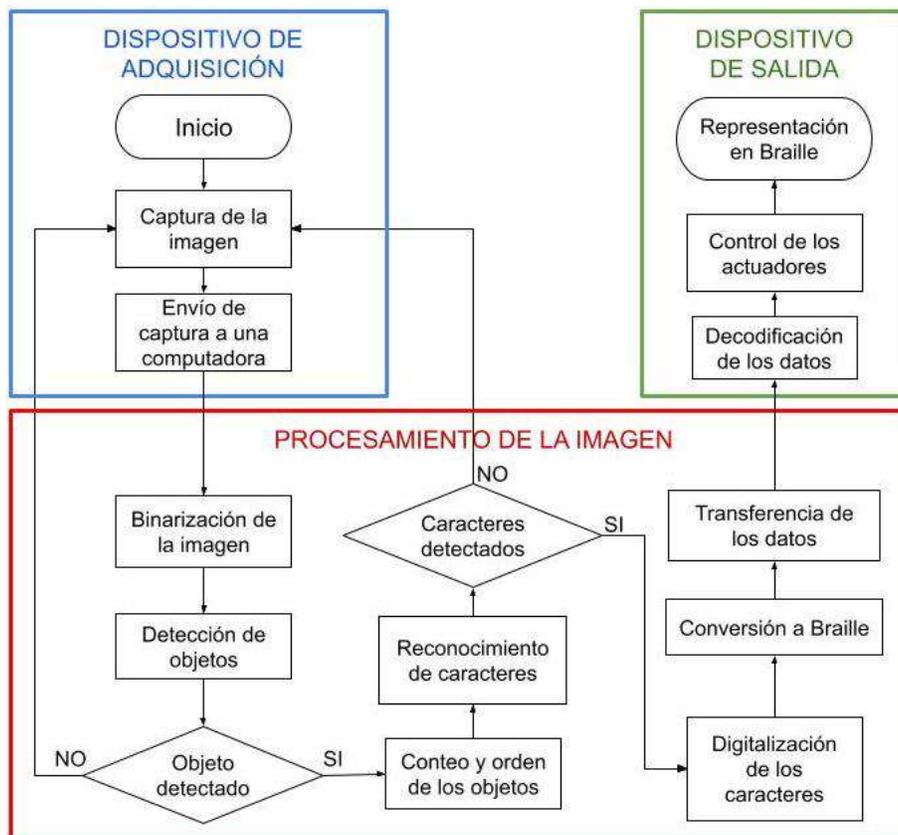


Figura 3-3. Diagrama metodológico extendido (Autoría propia, 2024).

La visualización más a detalle de la metodología se puede observar en la Figura 3-4, en esta figura se muestra los pasos que el sistema sigue para convertir el texto impreso en signos braille, y representarlos en una celda braille electrónica.

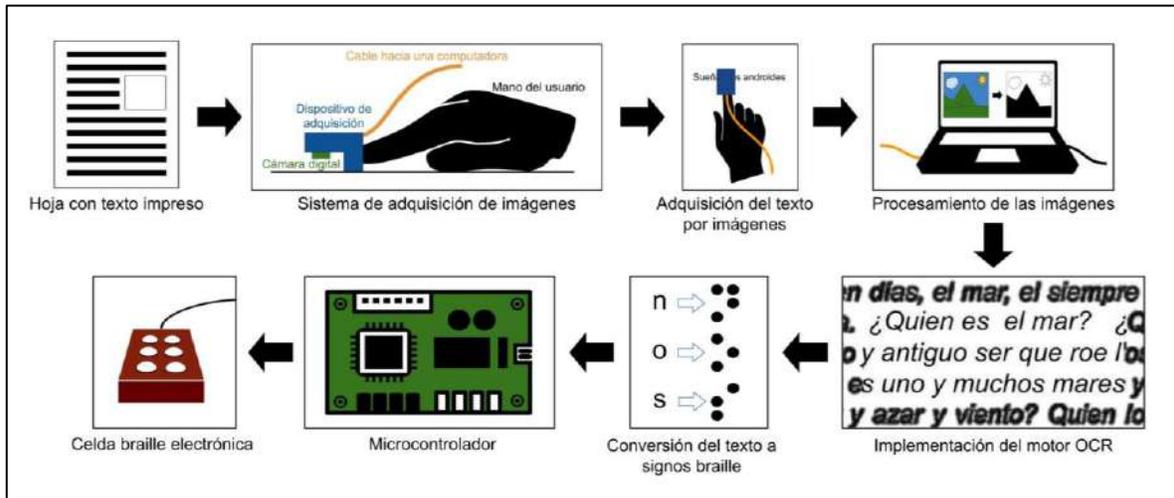


Figura 3-4. Diagrama detallado de la metodología para el sistema de traducción de texto impreso en español a signos braille (Autoría propia, 2024).

El primer paso del sistema consiste en tener una hoja con texto impreso con un fondo blanco. En el siguiente paso, es colocar el dispositivo de adquisición sobre la hoja y siempre debe estar conectado a la computadora, el usuario debe sostener el dispositivo de tal forma que la cámara digital apunte sobre el texto impreso. El tercer paso, es realizar movimientos horizontales sobre la hoja, para que la cámara capture el texto de un renglón y envíe las imágenes a la computadora. El cuarto paso, consiste en que una aplicación en la computadora procese las imágenes, esto con el fin de mejorar la calidad del texto u objetos que se encuentren en esta. En el quinto paso, es utilizar un motor OCR para identificar caracteres de texto en la fotografía. El siguiente paso, es convertir a signos braille todo aquel carácter detectado en la imagen. El séptimo paso, es simplemente enviar los datos de los signos braille a un microcontrolador, este último convertirá la información recibida en señales de salida para activar o desactivar reveladores, los cuales alimentan a unos solenoides en la celda braille. El último paso es alimentar con energía eléctrica un circuito de potencia, el cual suministrar voltaje y corriente a los solenoides que se encuentran en la celda braille; dependiendo de las señales de salida en el microcontrolador el circuito de potencia alimentara a los solenoides, y estos últimos moverán los émbolos para formar un signo braille.

La metodología explicada en el párrafo anterior es un proceso que el sistema efectuará continuamente, esto con el objetivo de que la información se actualice en el programa de la computadora y que el signo braille cambie acorde al texto capturado en la nueva imagen. La etapa de adquisición, el procesamiento de imágenes y del dispositivo de salida se describen en los siguientes subcapítulos; esto incluye los componentes y procesos que conforman a cada etapa.

3.3 Dispositivo de adquisición

El dispositivo de adquisición está compuesto por tres partes: una cámara digital, una base para esta misma y por una iluminación para hacer capturas de imágenes con una mejor calidad controlada, esto último con el fin de que las fotografías estén lo suficientemente claras y los detalles que contienen sean fácilmente reconocibles.

3.3.1 Cámara digital

La adquisición de las imágenes se realiza mediante el uso de una cámara digital dispuesta en un módulo, este mismo está conectado a la computadora por medio de un cable USB. En la Figura 3-5 se observa el módulo de la cámara que se utiliza en este dispositivo.



Figura 3-5. Módulo de cámara HBVCAM-W202012HD V33 (Autoría propia, 2024).

Las especificaciones del módulo de la cámara son la siguientes:

- Sensor: OV9726 CMOS (1/6 pulgada)
- Tasa de cuadros: 30 FPS
- Resolución: 1280 x 800
- Matiz: Color
- Píxeles: 1 MP
- Campo de visión: 42° - 70°
- Distancia focal: 3 mm - 2.8 mm
- Formato: MJPG / YUY2
- Salida: USB 2.0
- Enfoque: Manual (lente giratoria)
- Temperatura de operación: -30 a 70 °C
- Tamaño: 30 x 25 x 11.3 mm

3.3.2 Control de iluminación

El módulo de cámara mostrado en la Figura 3-5 tiene la capacidad de adaptarle un control de auto iluminación por medio de una foto resistencia y luces led, esto con el fin de que siempre existan capturas de imágenes con una misma iluminación o parecida a la luz natural. Los componentes que son utilizados para este propósito se muestran en la Figura 3-6, el inciso (a) en la imagen es una fotorresistencia, esta mide la intensidad de iluminación en el entorno; mientras que el inciso (b) es un led de luz blanca, el cual se ilumina con una cierta intensidad y esto último es controlado por la fotorresistencia.

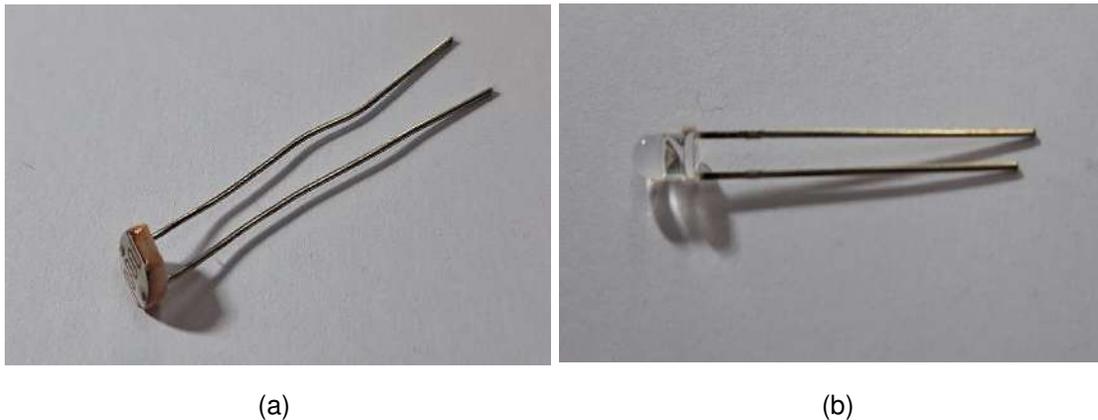


Figura 3-6. Componentes de control de iluminación de la cámara: (a) fotorresistencia, (b) led de luz blanca (Autoría propia, 2024).

En la parte frontal del módulo de la cámara digital se soldarán cuatro leds de luz blanca y una fotorresistencia, en la Figura 3-7 se muestra dicho resultado. La fotorresistencia es utilizada para detectar la cantidad de intensidad luminosa que llega al módulo, y dependiendo de la cantidad de luz en el entorno esta misma activará los leds con una cierta intensidad.

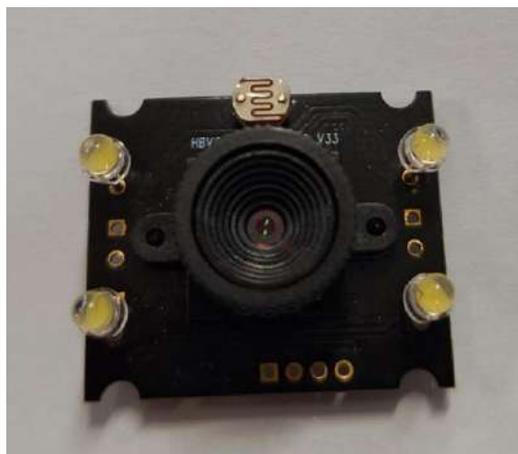


Figura 3-7. Módulo HBVCAM-W202012HD con auto iluminación led (Autoría propia, 2024).

3.3.3 Base para el módulo de la cámara

La cámara del dispositivo de adquisición requiere estar a una cierta distancia de la superficie con el texto impreso, y obtener un mejor detalle en las imágenes, por esto mismo el módulo de la cámara requiere una base en donde será montado, y para que la cámara siempre esté a una distancia óptima de captura, lo cual también se debe al enfoque manual de la cámara. Para cumplir con el propósito antes descrito, se contruyó una base para el módulo, esta fue diseñada en un programa CAD (Diseño Asistido por Computadora, *Computer Aided Design*) y fabricada por medio de la manufactura aditiva o impresión 3D, el material usado es PLA de color blanco para la base, tal y como se muestra en el inciso (a) de la Figura 3-8, también se diseñó y fabricó una tapa para el módulo de la cámara, esta es mostrada en el inciso (b) y la cual está hecha del mismo material, pero de color negro.

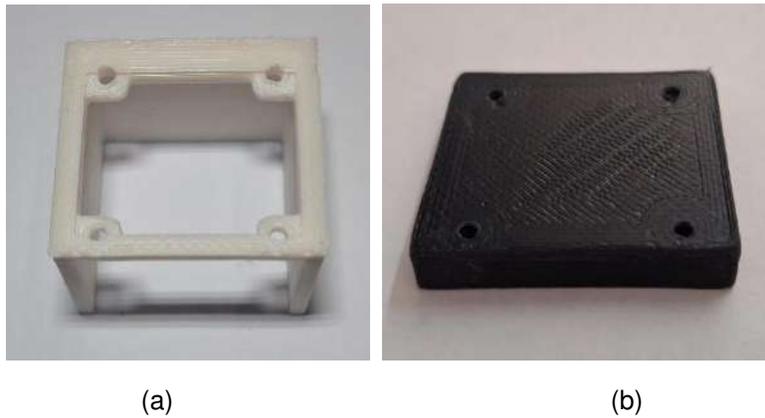


Figura 3-8. Partes del dispositivo de salida: (a) base para el módulo de la cámara, (b) tapa para el módulo de la cámara (Autoría propia, 2024).

En conjunto con las dos piezas anteriores, también se diseñó y fabricó, con los mismos métodos, una pequeña pieza, la cual sirve como una especie de tapa para la fotorresistencia, pero con una abertura, esto con el propósito de que la fotorresistencia reciba luz desde una sola dirección, y que preferentemente esta luz provenga del reflejo en la hoja con el texto impreso. La Figura 3-9 muestra esta pequeña pieza.



Figura 3-9. Tapa de la fotorresistencia (Autoría propia, 2024).

3.3.4 Pruebas del dispositivo de adquisición

El dispositivo de adquisición acoplado con las tres partes se puede observar en la Figura 3-10, en estas dos imágenes, se muestra, que el módulo de la cámara esta atornillado en conjunto con la base y su tapa; otro aspecto que se puede ver en la imagen del inciso (a) es la pequeña tapa que cubre una zona, que es donde se encuentra la fotorresistencia, esto es acorde a lo que se mencionó previamente sobre dicha pieza y la función que cumple esta misma. En cuanto a cómo debe ser usado el dispositivo de adquisición, la imagen que se muestra en el inciso (b) ejemplifica la disposición de este sobre una superficie plana, es decir, la cámara siempre tiene que estar apuntando hacia la superficie donde se encuentre el texto, y dependiendo de la iluminación que éste dentro del campo de visión de la cámara, los leds iluminarán la zona con una cierta intensidad.



Figura 3-10. Módulo de la cámara montada en la base con una tapa en la fotorresistencia: (a) vista inferior, (b) vista frontal (Autoría propia, 2024).

Otro aspecto más a considerar es la conexión de este dispositivo con la computadora, este siempre debe estar conectado con esta última, para ello un cable USB 2.0 se une al módulo de la cámara por la parte posterior del dispositivo, la Figura 3-11 muestra dicho cable.



Figura 3-11. Módulo de la cámara en la base con la tapa conectado a la computadora, vista superior (Autoría propia, 2024).

El funcionamiento del dispositivo se observa en las siguientes dos imágenes, la Figura 3-12 y Figura 3-13, siendo la primera una captura con iluminación externa, es decir, que la fuente de luz proviene de fuera del dispositivo y la segunda fotografía es tomada con la luz proporcionada por los leds acoplados al módulo. En la primera imagen se observa con claridad los detalles que contiene, las formas de las letras, las imperfecciones del papel y que gran parte de la imagen esta iluminada por la fuente externa; en cuanto a la imagen de la Figura 3-13, los detalles de las letras se conservan, pero las imperfecciones del papel se pierden y la iluminación no es consistente. Lo más importante en las capturas son los detalles de las letras, y que la iluminación sea muy similar en ambas, por ende, el control de iluminación que tiene el módulo es adecuado para hacer capturas de imágenes; además el procesamiento que se realiza posteriormente considera principalmente los detalles en los objetos de la imagen, es decir, letras, símbolos y otras formas con colores que contrasten con el fondo blanco de la hoja de papel.

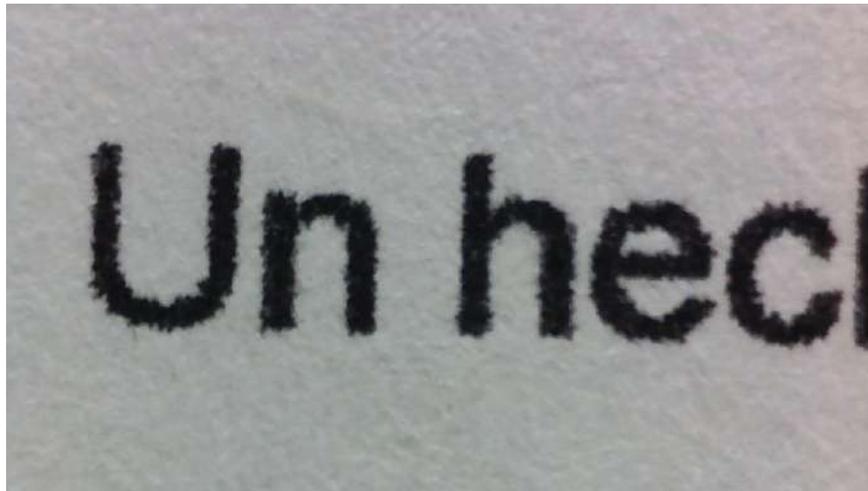


Figura 3-12. Captura de la cámara digital con iluminación externa (Autoría propia, 2024).

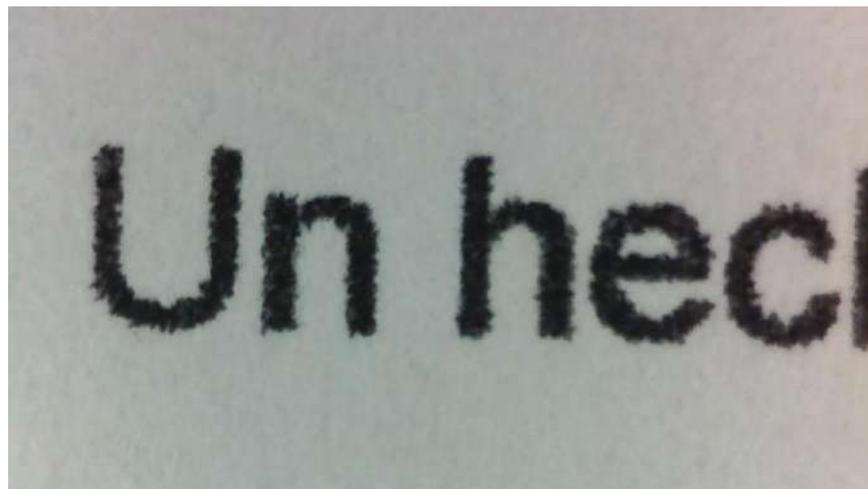


Figura 3-13. Captura de la cámara digital con la auto iluminación (Autoría propia, 2024).

3.4 Procesamiento de imágenes

La etapa de procesamiento de imágenes está compuesta por un programa de computadora, en el cual se implementan una serie de métodos para el procesamiento de imágenes y algoritmos, con el objetivo de extraer determinadas características de la imagen, para posteriormente realizar distintas operaciones en ella, y mejorar la imagen que es utilizada en el motor OCR. Las partes que conforman esta etapa del sistema de traducción es solamente el programa de computadora, por esto mismo lo que se explicará en esta sección son los métodos implementados en el procesamiento de las imágenes capturadas, la forma en cómo funciona el programa de computadora, y las pruebas que se realizan con los métodos de procesamiento para hacer operaciones en la imagen.

3.4.1 Métodos de procesamiento de imágenes

Los métodos de procesamiento de imágenes que se utilizan en el programa de computadora trabajan en una serie de pasos para llegar a un objetivo, es decir, que la implementación de los métodos inicia con la imagen enviada por la cámara digital, para que posteriormente cada uno estos se ejecuten uno detrás del otro, hasta obtener la imagen deseada o las características de esta. Los distintos métodos de procesamiento utilizados en el programa se describen a continuación:

- **Imagen a color**

La imagen a color es la captura que la cámara digital transfiere al programa de computadora, como la mostrada en la Figura 3-14. La ecuación (7) es la expresión matemática de la fotografía a color, donde x y y son las coordenadas de los pixeles; en cuanto a la intensidad de color en cada píxel, la ecuación (8) expresa el valor de colores rojo, verde y azul respectivamente. Estas dos ecuaciones, resumen lo que es una imagen digital a color, de tal forma que permite realizar operaciones y transformaciones a esta misma.

$$f(x, y) = \begin{pmatrix} f(0,0) & f(0,1) & \dots & f(0, N - 1) \\ f(1,0) & f(1,1) & \dots & f(1, N - 1) \\ \vdots & \vdots & \dots & \vdots \\ f(M - 1,0) & f(M - 1,1) & \dots & f(M - 1, N - 1) \end{pmatrix} \quad (7)$$

$$f(x, y) = \begin{bmatrix} f_R(x, y) \\ f_G(x, y) \\ f_B(x, y) \end{bmatrix} \quad (8)$$



Figura 3-14. Imagen a color (Autoría propia, 2024).

- **Escala de grises**

La conversión de una imagen a color en escala de grises se hace mediante el uso de la ecuación (9). En ésta los valores de intensidad del píxel en el color rojo, verde y azul se multiplican con unos coeficientes determinados para cada intensidad, esto se realiza para obtener del píxel un valor de intensidad que se corresponda en la escala de grises. Esta transformación se puede observar en la Figura 3-15, la cual es el resultado de la conversión a escala de grises de la imagen a color mostrada en la Figura 3-14.

$$g(x, y) = 0.299f_R(x, y) + 0.587f_G(x, y) + 0.114f_B(x, y) \quad (9)$$

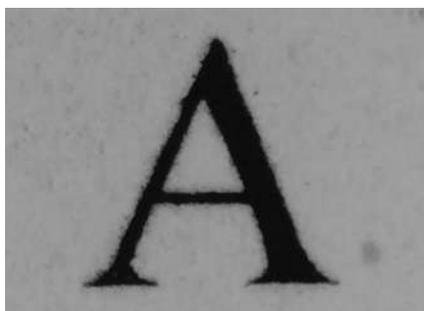


Figura 3-15. Imagen en escala de grises (Autoría propia, 2024).

- **Desenfoque gaussiano**

Este método de procesamiento es utilizado para difuminar imágenes, en este caso las fotografías que se encuentran en escala de grises. Este proceso genera en la imagen un efecto de desenfoque, el cual elimina el ruido en esta misma permitiendo tener objetos con contornos menos puntiagudos y con una forma más suave. El efecto que genera este proceso se muestra en la Figura 3-16, la cual es el resultado de aplicar el desenfoque gaussiano a la imagen de la Figura 3-15, y como se observa la imagen esta borrosa, pero el objeto contenido en esta

misma tiene bordes con esquinas más circulares y menos puntiagudas. El desenfoque gaussiano para las dos dimensiones es ejecutado por la ecuación (10), en la cual x y y son las distancias desde el origen en el eje horizontal y vertical respectivamente; y los valores de σ_x y σ_y son la desviación estándar en la dirección de 'x' y 'y' respectivamente.

$$G(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)} \quad (10)$$

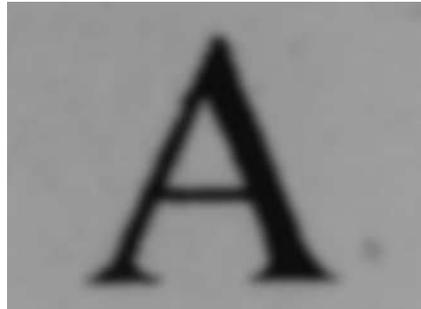


Figura 3-16. Imagen en escala de grises con el desenfoque gaussiano (autoría propia, 2023).

- **Imagen en blanco y negro**

La transformación de una imagen en escala de grises a blanco y negro se lleva a cabo por la ecuación (11), en esta el x y y son las coordenadas de un píxel de una imagen en escala de grises, y *umbral* es un valor para determinar la conversión a blanco y negro del píxel seleccionado, es decir, dependiendo de si el valor en la intensidad en gris del píxel es mayor o menor al *umbral*, este mismo obtiene un nuevo valor en la intensidad de 0 o 1, los cuales son el color negro y blanco respectivamente. Esta transformación se puede ver en la Figura 3-17 como el resultado de convertir la imagen de la Figura 3-16 a blanco y negro.

$$h(x, y) = \begin{cases} 1 & \text{umbral} \geq g(x, y) \\ 0 & \text{umbral} < g(x, y) \end{cases} \quad (11)$$

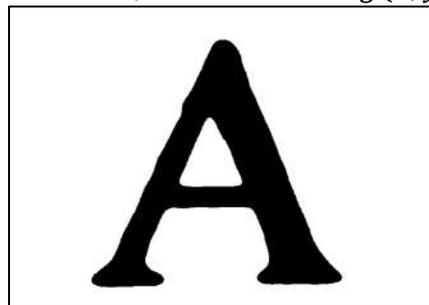


Figura 3-17. Imagen en blanco y negro (Autoría propia, 2024).

- **Algoritmo de Canny**

El algoritmo Canny es un proceso que permite detectar los bordes de objetos en imágenes. En este caso es utilizado para encontrar los contornos de imágenes en blanco y negro, como la mostrada en la Figura 3-17, el resultado que arroja este algoritmo para la imagen previa se puede observar en la Figura 3-18, en esta se ven líneas de color blanco, las cuales forman los bordes que separan los pixeles de color negro con los de blanco. El desenfoco gaussiano utilizado previamente permite al algoritmo de Canny, detectar contornos mejor interconectados, es decir, que la línea que separa los pixeles blancos y negros este mejor conectada entre sí, para que de esta manera se formen los contornos conforme al objeto contenido.



Figura 3-18. Imagen con la implementación del algoritmo Canny (Autoría propia, 2024).

- **Algoritmo rectángulo delimitador (Bounding Rectangle)**

Este proceso es implementado para contener dentro de un rectángulo un contorno, en este caso el borde exterior de los objetos contenidos en la imagen. El rectángulo formado contiene ciertas características: tamaño del rectángulo, el ancho y alto; la coordenada de la esquina superior izquierda; la coordenada de la esquina inferior derecha; el ángulo de inclinación del rectángulo; el área del rectángulo y el centro de este mismo.

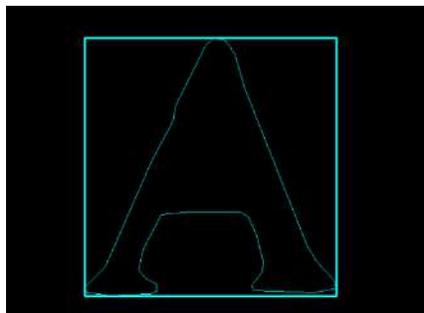


Figura 3-19. Imagen con la implementación de algoritmo rectángulo delimitador (*Bounding Rectangle*) (Autoría propia, 2024).

- **Rotación de la imagen**

Este es un método utilizado para rotar la imagen desde un centro de rotación especificado, para realizar este proceso se implementa la ecuación (12), donde los valores de x y y son las coordenadas del píxel a rotar, θ es el ángulo de rotación, x_c y y_c son las coordenadas del centro de rotación; por último, x' y y' son las nuevas coordenadas del píxel rotado. El ejemplo que se observa en la Figura 3-20 es la imagen a color que pertenece a la Figura 3-14, la cual esta rotada -30° con respecto a su centro.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & x_c(1 - \cos\theta) + y_c \sin\theta \\ \sin\theta & \cos\theta & y_c(1 - \cos\theta) - x_c \sin\theta \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (12)$$



Figura 3-20. Imagen a color rotada (Autoría propia, 2023).

3.4.2 Programa de computadora

El programa de computadora es el medio por el cual las imágenes capturadas por la cámara son procesadas, además, en este se implementan los métodos de procesamiento mencionados en la sección anterior, tal aplicación de computadora se puede observar en la Figura 3-21. En este programa la imagen a color tiene una secuencia de procesamiento: primero se convierte a escala de grises, después se aplica el desenfoque gaussiano, el tercer paso es convertirla a blanco y negro, el cuarto paso es obtener los bordes de los objetos con el algoritmo Canny, para que después se obtengan las características por medio de rectángulo delimitador, y dependiendo del ángulo y posición del objeto se aplica a la imagen el proceso de rotación; por último, en la imagen rotada, se vuelve a usar el algoritmo Canny y el método de rectángulo delimitador para obtener las nuevas posiciones de los objetos en la imagen, además de seleccionar y recortar la fotografía del objeto que debe entrar en el OCR. En cuanto al uso del motor OCR, el programa utiliza el Tesseract OCR, éste implementa una red neuronal de la forma LSTM para identificar diferentes tipos de caracteres, símbolos, números y

palabras; dependiendo del objeto seleccionado en la imagen, el OCR lo detectará y almacenará como un carácter en código ASCII.

El procesamiento para una sola imagen se detiene cuando el objeto es seleccionado, recortado y enviado al OCR para identificar un carácter, símbolo o número, pero una vez detectado y almacenado en código ASCII, éste último, le prosigue una conversión al signo braille español acorde a las normas definidas por la CBE; una vez guardado el signo braille, este es enviado por el programa en forma de un byte a una tarjeta con un microcontrolador, para que esta última controle una celda electrónica que representará físicamente el signo braille. El procedimiento antes mencionado en conjunto con los métodos de procesamiento de imágenes es la forma en cómo funciona el programa de computadora, pero este debe estar trabajando de manera continua, es decir, tiene que capturar la fotografía, procesar la imagen, detectar el símbolo o carácter, convertir a signo braille y enviar los datos a la tarjeta, y debe repetir este proceso incesantemente hasta que se le ordene detenerse.

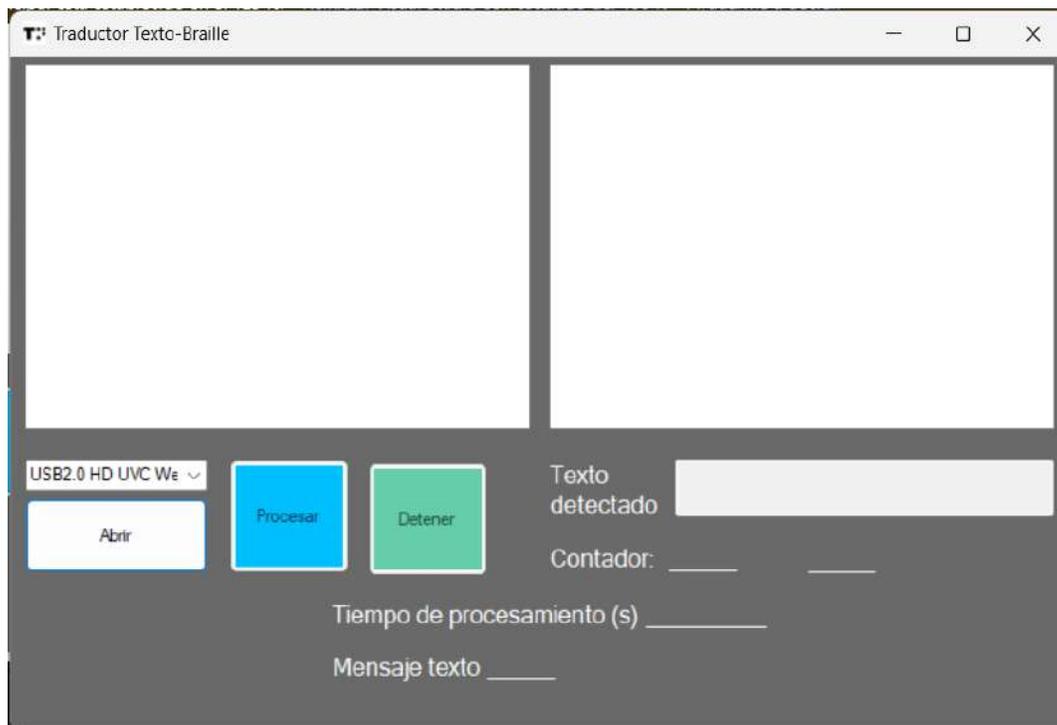


Figura 3-21. Programa de computadora para el procesamiento de imágenes (Autoría propia, 2024).

La función de cada una de las partes en el programa de computadora se muestra en la Figura 3-22, en esta se pueden observar dos grandes cuadros blancos, el de la izquierda tiene la función de mostrar las imágenes que captura la cámara, en forma de video, pero sin grabar o guardar las imágenes; el cuadro de la derecha es el lugar donde se presenta la imagen completamente procesada. Las otras partes de la aplicación se encuentran debajo de los dos cuadros

grandes, el primero a la izquierda es un menú de selección de cámaras digitales, en el cual se debe elegir una cámara y debajo está el botón que habilita la cámara para capturar imágenes; en el funcionamiento del programa, es necesario siempre abrir la cámara antes de realizar el procesamiento de imágenes. Los botones azul y verde tiene la finalidad de iniciar y detener respectivamente el procesamiento de las imágenes, es decir, que el botón azul da comienzo con la captura de la imagen, el procesamiento que ésta recibe por parte de los diferentes métodos, la identificación del carácter en el OCR, la conversión a signo braille y el envío de los datos a la tarjeta; esto lo realiza cíclicamente hasta que se presiona el botón verde, el cual tiene propósito de detener todo el proceso.

Las demás partes del programa tienen la finalidad de mostrar los datos referentes al proceso de traducción de texto impreso a braille, en este apartado se encuentra el pequeño rectángulo blanco a la derecha, este tiene la finalidad de mostrar los caracteres que el OCR identifica; debajo de este rectángulo está un contador, el cual enumera las veces que el mismo carácter es detectado, y a la derecha de este último está una línea, donde se despliega un mensaje relacionado con el signo braille enviado a la tarjeta, es decir, si el signo braille es simple, este se representará en un solo carácter, y si el un signo es complejo, el carácter a mostrar es un número, una letra mayúscula u otro que requiera dos signos braille. Entre las últimas dos partes está un medidor de tiempo, éste indica los segundos que tarda todo el proceso desde la captura de una imagen hasta que se envían los datos a la tarjeta, este contador se reinicia cada vez que se vuelve a capturar una fotografía, procesarla y enviar los datos. Por último, está una línea de mensaje relacionada con el cuadro donde se muestra la imagen procesada, la función de este es desplegar un mensaje, el cual indica si hay un objeto dentro de la imagen.

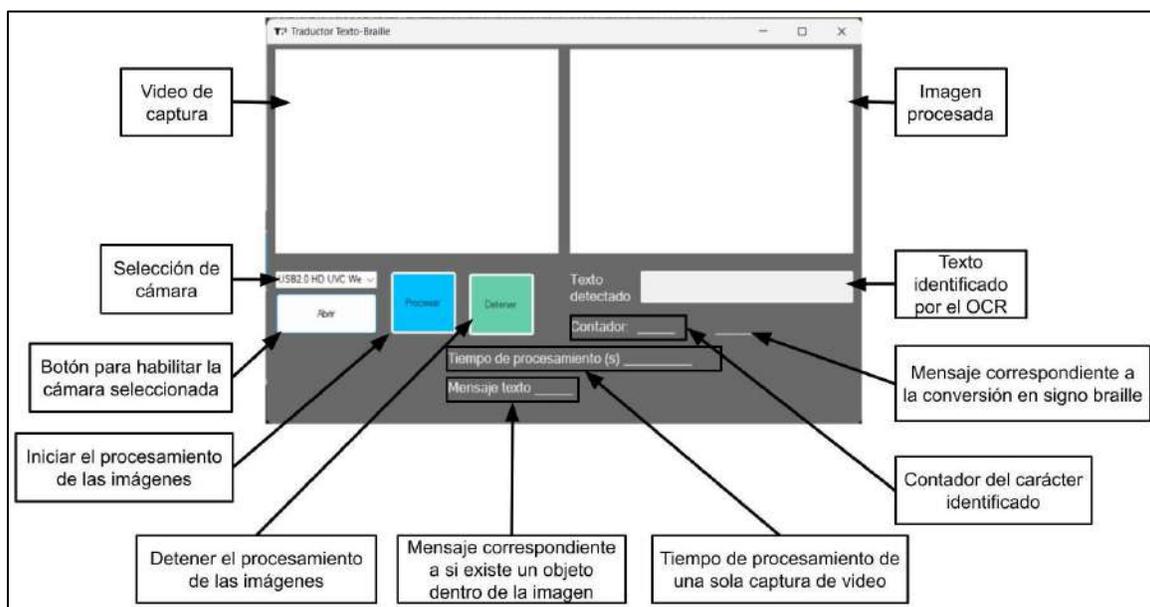


Figura 3-22. Función de cada parte del programa de computadora para el procesamiento de imágenes (Autoría propia, 2024).

Algo más a mencionar son las características del sistema computacional que es utilizado por el programa de computadora, las cuales son las siguientes:

- Sistema operativo Windows 11.
- Sistema de 64 bits x64.
- Procesador de 8 núcleos.
- Velocidad base del procesador 3.2 GHz.
- Memoria RAM de 16 GB.
- Tarjeta gráfica de 6 GB de VRAM.
- Frecuencia base de tarjeta gráfica 817 MHz

3.4.3 Pruebas del procesamiento de imágenes

En esta sección se explica el uso de los diferentes métodos de procesamiento de imágenes, enfocados a dos procesos que se implementa en el programa de computadora. El primero consiste en detectar el ángulo de inclinación del texto, para después usar este dato en el método de rotación de imagen, y de esta forma corregir la orientación del texto el cual debe estar de manera horizontal; en cuanto al segundo proceso, este es una técnica para seleccionar un carácter en el texto horizontal, para ello se utiliza la gran mayoría de los métodos de procesamiento de imágenes, además de desarrollar algoritmos que seleccionen de una buena manera el objeto, el cual es usado por el OCR para detectar un carácter.

❖ Cálculo del ángulo

Este primer proceso consiste en principalmente en hacer rotar una imagen de manera automática, para lograr este objetivo primero se detecta el ángulo de inclinación del texto y después se efectúa el método de rotación en la imagen, pero es necesario pasar por una serie de pasos intermedios que permiten obtener este dato de una manera conveniente, cada uno de esto son explicados a continuación.

1. El primer paso es la captura de una imagen de color, esta se muestra en la Figura 3-23, y en ella se puede observar que el texto tiene un ángulo con respecto a la horizontal.

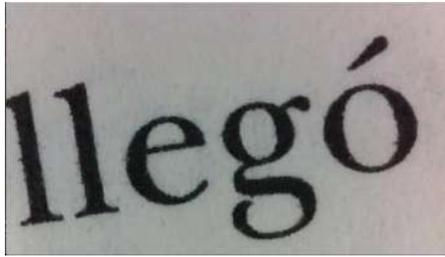


Figura 3-23. Captura de imagen a color con texto inclinado (Autoría propia, 2024).

2. El siguiente paso es convertir la imagen de color capturada en escala de grises, el resultado es mostrado en la Figura 3-24.

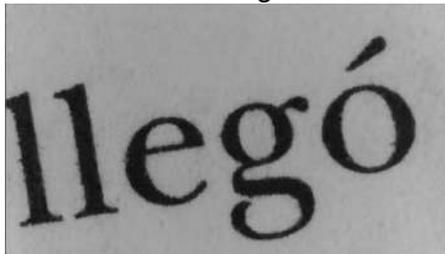


Figura 3-24. Imagen en escala de grises con texto inclinado (Autoría propia, 2024).

3. El tercer paso es aplicar el desenfoque gaussiano en la imagen en gris, tal y como se puede ver en la Figura 3-25.



Figura 3-25. Imagen con desenfoque gaussiano del texto inclinado (Autoría propia, 2024).

4. Después la imagen con desenfoque gaussiano es convertida a blanco y negro, esto se observa en la Figura 3-26.



Figura 3-26. Imagen en blanco y negro del texto inclinado (Autoría propia, 2024).

5. El quinto paso es aplicar el algoritmo Canny para identificar los bordes de los objetos, en este caso las formas de las letras que están en la imagen, la Figura 3-27 muestra el resultado.



Figura 3-27. Algoritmo Canny aplicado en la imagen del texto inclinado (Autoría propia, 2024).

6. El siguiente paso es usar algoritmo del rectángulo delimitador, este se aplica en la imagen con los bordes de las letras para obtener rectángulos que encapsulan los contornos de las letras, la respuesta generada por este método se muestra en la Figura 3-28.



Figura 3-28. Algoritmo del rectángulo delimitador implementado en la imagen del texto inclinado (Autoría propia, 2024).

7. El séptimo paso es crear un algoritmo que determine cuáles son los objetos que se encuentran dentro de la imagen, además de que no estén en contacto con el borde de esta misma. La característica utilizada para este propósito es la posición de las esquinas, estas forman los rectángulos que contienen a cada uno los objetos mostrados en la Figura 3-28. El resultado de este proceso se observa en la Figura 3-29, donde los recuadros en verde son los contornos que encapsulan a cada objeto, mientras que los de color rojo, muestran la posición de las esquinas de los rectángulos verdes, y cómo, estas últimas, no tocan el borde de la imagen.

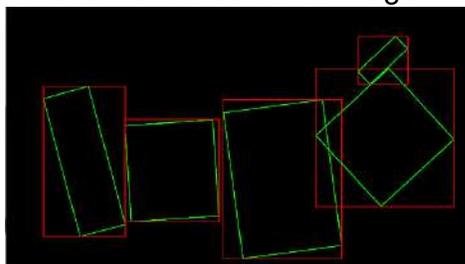


Figura 3-29. Objetos de la imagen dentro de la zona de interés (Autoría propia, 2024).

8. El paso que sigue es determinar con un algoritmo qué objeto tiene el área más grande A_{max} , además cuáles objetos tienen un porcentaje en área mayor a 60% con respecto al objeto más grande, para esto último se

utilizan las ecuaciones (13) y (14); donde la primera calcula el porcentaje del área de un objeto p_i , y la segunda determina si este mismo cumple con la condición del 60%, dependiendo de cumplir con esta restricción, las características del objeto se guardan en un vector y el N_o (número de objetos) aumenta; N_A es el número de áreas y, A_i es el área del objeto usada en las ecuaciones. El resultado de este proceso se puede ver en los recuadros azules que se muestran en la Figura 3-30.

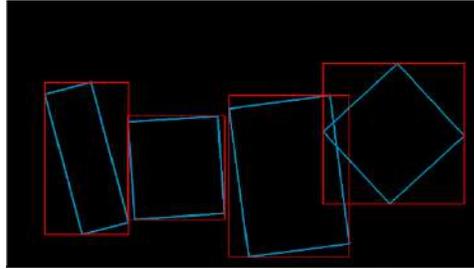


Figura 3-30. Objeto de la imagen con la mayor área (Autoría propia, 2024).

$$p_i = \frac{A_{max} - A_i}{A_{max}} 100 \quad (13)$$

$$N_o = \sum_{i=0}^{N_A} \begin{cases} 1 & p_i \geq 60\%A_{max} \\ 0 & p_i < 60\%A_{max} \end{cases} \quad (14)$$

9. El noveno paso es obtener el ángulo de inclinación del texto, para esto se calcula, de cada recuadro, su ángulo, además de que la separación entre un ángulo no sea mayor a 20° con respecto a los otros; es decir compara un ángulo con respecto a un total de N_o ángulos, si su diferencia con respecto a otro ángulo es mayor a 20° , entonces un valor N no aumenta, en caso de lo contrario N se incrementa y, las características del vector se guardan en un vector distinto. Para obtener el ángulo de rotación se utiliza la ecuación (15), θ_k es el ángulo de inclinación de un rectángulo, θ_r es el ángulo de rotación que se calcula y N es el número de rectángulos utilizados para calcular la rotación, la Figura 3-31 muestra los objetos usados en el cálculo del ángulo de rotación.

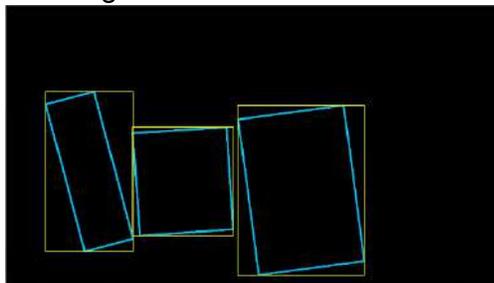


Figura 3-31. Recuadros usados en el cálculo del ángulo de rotación (Autoría propia, 2024).

$$\theta_r = \frac{\sum_{k=1}^N \theta_k}{N} \quad (15)$$

10. El último paso es aplicar la transformada de rotación con el ángulo calculado, el resultado de este método se observa en la Figura 3-32, la cual, es la imagen que se ve en la Figura 3-26, pero rotada un ángulo que permite mostrar al texto de manera horizontal. La transformación se puede comprobar por la letra 'l' más a la izquierda de la imagen, debido a que en la punta superior de esta letra falta un cierto pedazo, el cual, si uno observa la imagen original, es la parte que se encuentra en contacto con el borde.



Figura 3-32. Imagen en blanco y negro rotada un cierto ángulo (Autoría propia, 2024).

❖ Selección de carácter

El segundo proceso es la selección de un carácter de manera automática, esta técnica es un conjunto de pasos en los cuales se eligen qué objeto debe entrar en el OCR, para que éste último identifique el carácter y lo almacene en código ASCII. La serie de pasos que se describen a continuación es una forma general de entender este proceso de selección.

1. El primer paso es tener una imagen con texto en horizontal, tal y como se muestra en la Figura 3-33. Esta fotografía, tiene el aspecto de haber pasado por los procesos de captura a color, la conversión a escala de grises y el desenfoque gaussiano; es por eso por lo que se encuentra en blanco y negro, y que el texto se encuentra de manera horizontal.



Figura 3-33. Imagen en blanco y negro con texto horizontal (Autoría propia, 2024).

2. El segundo paso es aplicar el algoritmo Canny a la imagen, esto genera la respuesta mostrada en la Figura 3-34, la cual es obtener los bordes de cada uno de los objetos dentro de la imagen.



Figura 3-34. Algoritmo Canny aplicado a la imagen de blanco y negro con texto horizontal (Autoría propia, 2024).

3. El tercer paso es implementar el algoritmo del rectángulo delimitador, el cual construye los recuadros que contiene a todos los objetos que formaron un borde en la imagen, esto se puede ver en la Figura 3-35.



Figura 3-35. Algoritmo de rectángulo delimitador implementado en la imagen con texto horizontal (Autoría propia, 2024).

4. El siguiente paso es determinar los objetos que se encuentran dentro de la imagen, además de que sus contornos no estén en contacto con algún borde. Esto mismo se logra por el uso de la posición de cada lado en los recuadros, el resultado que se obtiene en este proceso para la imagen de la Figura 3-35, se muestra en la Figura 3-36, donde se puede apreciar solamente a los rectángulos que no están en contacto con el borde de la imagen.

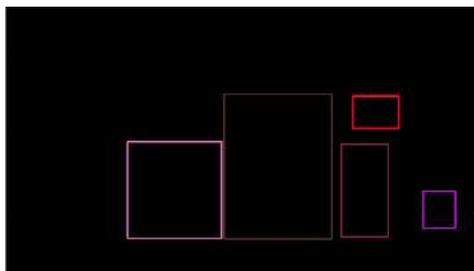


Figura 3-36. Recuadros dentro de la imagen (Autoría propia, 2024).

5. El quinto paso es crear un algoritmo que construya contornos con otros más pequeños, este proceso debe considerar a rectángulos que estén de

manera vertical, es decir, que el nuevo recuadro más grande solo se crea cuando uno está encima o debajo del otro, tal y como se ve en el cuadro rojo de la Figura 3-37. Este proceso también considera a los otros rectángulos que no tienen otro arriba o debajo ellos, pero a estos no les construye un nuevo recuadro, en la imagen, son los cuadros amarillos. El propósito de este paso es contener la letra 'i', en este caso, pero también es aplicado a letras o símbolos que estén compuestos de otras formas.

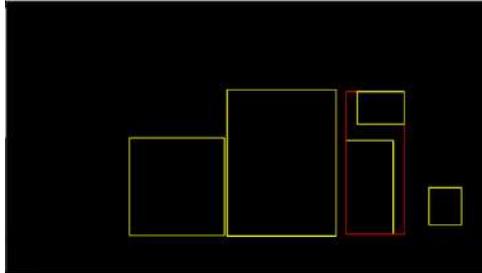


Figura 3-37. Construcción de bordes más grandes (Autoría propia, 2024).

6. El sexto paso consiste seleccionar los contornos más grandes de la imagen, es decir, forma una imagen con los recuadros que contenga a otros más pequeños; el resultado de este proceso para la imagen anterior se observa en la Figura 3-38.

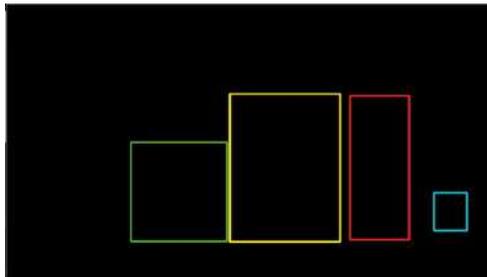


Figura 3-38. Recuadros más grandes de la imagen (Autoría propia, 2024).

7. El séptimo paso, es crear un algoritmo para seleccionar el cuadro más centrado en la imagen, es decir, elegir entre los rectángulos más grandes, el que se encuentre más cerca del centro de la imagen. La respuesta que se obtiene de este proceso de la Figura 3-38 se muestra en la Figura 3-39.

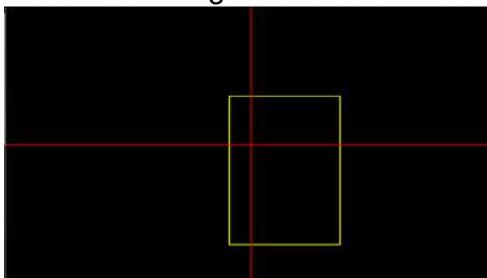


Figura 3-39. Recuadro grande más centrado en la imagen (Autoría propia, 2024).

8. El último paso es realizar un recorte en la imagen de blanco y negro del texto horizontal, esto mediante el uso de las dimensiones y coordenadas del recuadro más centrado. El resultado es mostrado en la Figura 3-40.

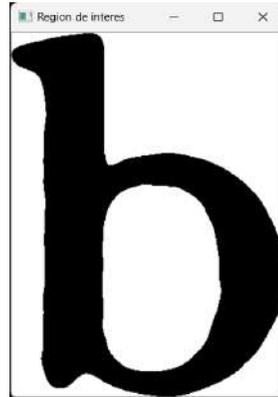


Figura 3-40. Sección recortada de la imagen en blanco y negro con texto horizontal (Autoría propia, 2024).

3.5 Dispositivo de salida

El dispositivo de salida está compuesto por tres partes: la primera es el sistema de comunicación y control, la segunda es la etapa de potencia, y la última es la parte de la representación física del signo braille. Estas tres partes en conjunto forman el bloque de salida de la metodología, y para que continuamente se den forma a los distintos signos braille transmitidos por el programa de computadora, son representados en la celda braille continuamente.

3.5.1 Comunicación y control

Esta parte del dispositivo de salida utiliza una tarjeta con un microcontrolador para la comunicación con la computadora, y también como el control de la etapa de potencia, la Figura 3-41 muestra la tarjeta con el microcontrolador.



Figura 3-41. Tarjeta Arduino UNO (Autoría propia, 2024).

Las especificaciones de la tarjeta son las siguientes:

- Microcontrolador: ATmega328P
- Voltaje de operación: 5V CD
- Voltaje de entrada: 7-12 V CD (recomendada)
- Pines de entradas y salidas digitales: 14
- Pines PWM: 6
- Pines de entrada analógica: 6
- Corriente DC por pines de entrada y salida: 20 mA
- Corriente DC para el pin 3.3 V: 50 mA
- Memoria Flash: 32 KB
- SRAM: 2 KB
- EEPROM: 1KB
- Frecuencia de reloj: 16 MHz
- Longitud: 68.6 mm
- Anchura: 53.4 mm

3.5.2 Etapa de potencia

En esta parte del dispositivo se encuentran los componentes electrónicos que permiten alimentar con voltaje y corriente a los solenoides, y por ende generar un movimiento en ellos. En la Figura 3-42 se muestra un circuito con diodos 1N4007 conectados a módulos con relevadores con el fin de evitar retorno de corrientes.

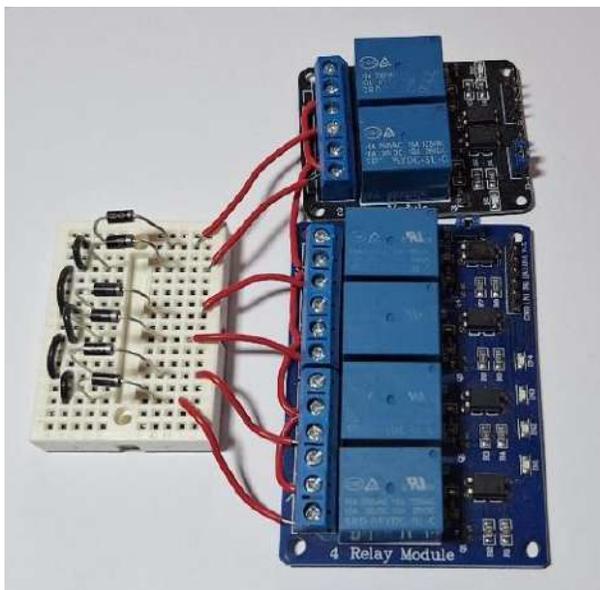


Figura 3-42. Etapa de potencia de la celda braille (Autoría propia, 2023).

Las especificaciones de los relevadores son las siguientes:

- Voltaje de alimentación: 5 V CD
- Corriente de disparo: 5 mA
- Carga AC máxima: 250 V / 125 V CD, 10 A
- Carga DC máxima: 30 V DC / 28 V CD, 10 A
- Corriente nominal de bobina: 75 mA
- Expectativa de vida: 100 000 accionamientos
- Temperatura de operación: -25 °C a 70 °C

3.5.3 Representación del signo braille

La representación física de un signo braille se realiza a través de la celda mostrada en la Figura 3-43, dentro de ésta se encuentran los solenoides. La base

para el arreglo de estos mismos fue diseñada en CAD y fabricada mediante manufactura aditiva, el material utilizado es PLA de color blanco.



Figura 3-43. Celda braille (Autoría propia, 2024).

El tipo de solenoide usado en la celda braille se muestra en la Figura 3-44, las especificaciones de estos mismos son las siguientes:

- Voltaje nominal: 12 V CD
- Corriente de carga: 200 mA
- Resistencia de bobina: 55 Ohm
- Recorrido del émbolo: 4 mm
- Tamaño: 2.9 x 1 x 1.1 cm
- Peso: 11.34 g



Figura 3-44. Electroimán solenoide (Autoría propia, 2024).

3.5.4 Pruebas del dispositivo de salida

El dispositivo de salida armado con las tres partes se muestra en la Figura 3-45, en esta se observa un cable proveniente de la computadora, el cual está conectado con la tarjeta del microcontrolador; también de la tarjeta se ven cables que salen hacia la etapa de potencia, la cual tiene 12V CD como el voltaje de entrada; los cables de color rojo alimentan al circuito de diodos y a los solenoides que están conectados en este mismo.

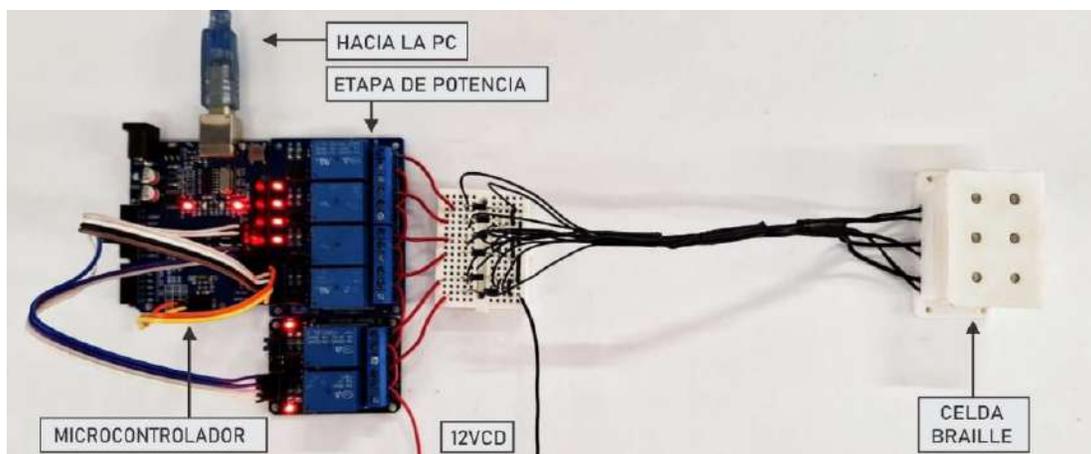


Figura 3-45. Dispositivos electrónicos y electromecánicos que componen la celda braille electrónica (Autoría propia, 2024).

El funcionamiento de este dispositivo consiste en lo siguiente: un dato es enviado desde la computadora en forma de byte, la tarjeta con el microcontrolador interpreta la información recibida y la convierte en señales eléctricas, estas mismas salen hacia los módulos de relevadores para activar la bobina de estos últimos, y por ende dejar pasar el voltaje de 12 V; dependiendo de qué relevador se accione, éste permitirá que un solenoide se alimente con el voltaje, al final el émbolo del electroimán se moverá y dará forma a un punto de signo braille. En este dispositivo se realizaron pruebas, en las que se enviaron los valores para representación de un signo braille correspondiente a un carácter de una letra, la Figura 3-46 muestra ejemplos de letras, que conforman palabras, representadas en la celda braille.



Figura 3-46. Conversión de caracteres de palabras (imagen izquierda) al sistema braille, y representados en la celda braille (imagen derecha) (Autoría propia, 2024).

El desarrollo del dispositivo de salida se utilizó en el artículo de nombre: *Dispositivo Electromecánico de Puntos en Relieve Actualizable para Reproducción de Caracteres del Sistema Braille*; este documento se encuentra anexado en el Apéndice A.

4 PRUEBAS Y RESULTADOS

El sistema de traducción de texto impreso a braille se sometió a pruebas, para comprobar el funcionamiento de este mismo y verificar su efectividad en la detección del texto impreso, el cual es convertido a signos braille. El sistema completo de traducción de texto impreso a braille se muestra en la Figura 4-1, donde se puede ver el dispositivo de adquisición dentro de un recuadro azul, este aparato está tomando capturas en una hoja de texto; otra parte del sistema en un recuadro rojo es el procesamiento de imágenes, se puede ver trabajando en la computadora, mirando de cerca en la pantalla, se muestra una letra, la cual es la fotografía que ha enviado el dispositivo de adquisición y el procesamiento que esta recibe; en un recuadro de color verde está el dispositivo de salida, encargado de recibir los datos del signo braille desde la computadora y de representarlo en la celda braille electrónica.

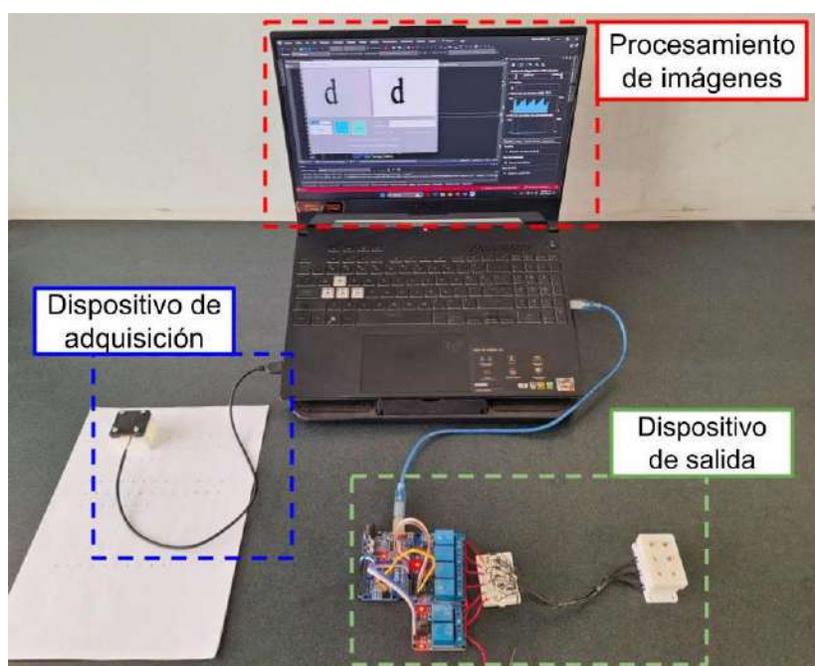


Figura 4-1. Sistema de traducción de texto al sistema braille (Autoría propia, 2024).

4.1 Pruebas del sistema

Las pruebas fueron hechas con hojas de texto impreso, la Tabla 1 especifica las condiciones del texto utilizado en las pruebas del sistema, en ella se usaron dos tipos de fuentes: Arial y Times New Roman, ambas con el mismo tamaño 12; el número de caracteres para las dos fuentes es de 94, esto es debido a que se realizaron capturas individuales en cada una de las letras del alfabeto, tanto minúsculas como mayúsculas, además de considerar las vocales con acento y diéresis, por lo que arrojaron una cantidad de 66 letras; también se consideraron en las pruebas los signos de puntuación: *el punto, la coma, los dos puntos, el*

punto y coma, los signos de interrogación de apertura y cierre, los signos de exclamación de apertura y cierre, las dobles comillas de apertura y cierre, las comillas angulares de apertura y cierre, el signo mayor que y menor que, y por último el signo guion; esto arroja una cantidad de 84 y sumando las diez cifras que van desde el cero hasta al nueve, el resultado es de 94 caracteres.

Tabla 1. Matriz de pruebas del sistema de traducción de texto impreso a braille.

Fuente	Tamaño de fuente	Número de caracteres	Capturas por carácter	Número de renglones	Interlineado de renglón
Arial	12	94	10	9	2.5
Times New Roman	12	94	10	9	2.5

En otro aspecto de las pruebas, está el número de capturas a los caracteres. El realizar varias tomas al mismo carácter es para corroborar la eficacia del OCR, y la correcta identificación de las letras, signo y números en la hoja de texto. La manera en cómo está distribuido el texto es de nueve renglones para ambas fuentes, esto se encuentra en la matriz de pruebas de la Tabla 1, pero también se pueden ver en los Apéndices B y C que son los textos en Arial y Times New Roman respectivamente, y los cuales son los textos utilizados en las pruebas del sistema de traducción de texto impreso a braille; también estuvo en consideración la separación de los renglones, que es de 2.5.

❖ **Secuencia de las pruebas**

Las pruebas que se realizaron fueron dos, una para el texto en Arial y otra en Times New Roman, pero ambas siguieron los siguientes pasos:

1. El primer paso es colocar el dispositivo de adquisición en la hoja de pruebas, al inicio del primer renglón.
2. El siguiente paso es hacer un movimiento horizontal hasta que la cámara este sobre un carácter de la hoja y éste se muestre en el programa de computadora, tal y como se observa en la Figura 4-2 para movimiento de la cámara, y en la Figura 4-3 para la imagen mostrada en el programa.
3. En el tercer paso la cámara envía las imágenes a la computadora y en esta el programa inicia el procesamiento de una sola imagen.
4. El cuarto paso consiste en esperar al programa de la computadora en el procesamiento de la imagen, así como la detección del carácter por parte del OCR.

5. Después el programa entrega el resultado del carácter identificado, este lo convierte en un signo braille, y por último envía los datos del signo braille a la tarjeta con el microcontrolador.
6. Este paso simplemente es observar el signo que se representa en la celda braille electrónica, y verificar si éste corresponde con el carácter detectado por el OCR, la Figura 4-4 muestra un ejemplo de la representación de un signo braille.
7. Este séptimo paso es similar al paso 2, pero simplemente es mover un poco el dispositivo de adquisición para que la cámara vuelva a capturar la letra en otra parte de la imagen.
8. En el octavo se vuelven a ejecutar los pasos desde el número 3 hasta el 6.
9. El paso número 7 y 8 se repiten otras ocho veces, que, desde el primer paso hasta el actual, son las diez capturas hechas a un mismo carácter.
10. El décimo paso es ejecutar nuevamente los pasos del número 2 al 9, esto se debe repetirse en todos los caracteres del renglón, y que estos mismos hayan sido procesados, identificados por el OCR y enviados a la tarjeta unas diez veces.
11. El undécimo paso es colocar el dispositivo de adquisición sobre un nuevo renglón.
12. Este paso consiste en repetir los pasos desde el 2 hasta el 11, esto debe realizarse hasta que todos los caracteres, de los nueve renglones se hayan terminado de procesar, identificar y representar en la celda electrónica.
13. El último paso es simplemente detener la captura y procesamiento de las imágenes, la identificación y la transmisión de datos a la tarjeta de la celda braille.

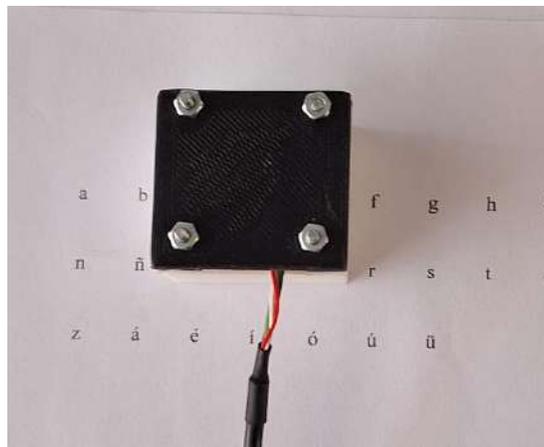


Figura 4-2. Dispositivo de adquisición sobre una hoja con texto impreso (Autoría propia, 2024).

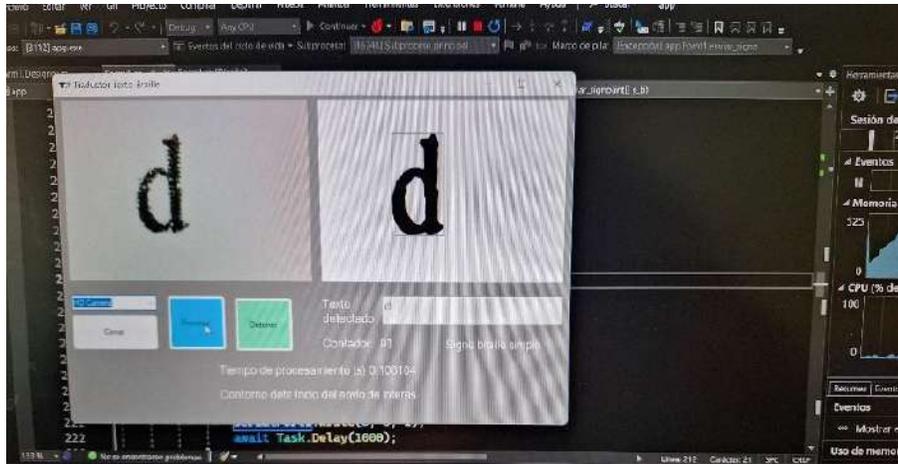


Figura 4-3. Programa de computadora realizando el procesamiento de imágenes (Autoría propia, 2024).

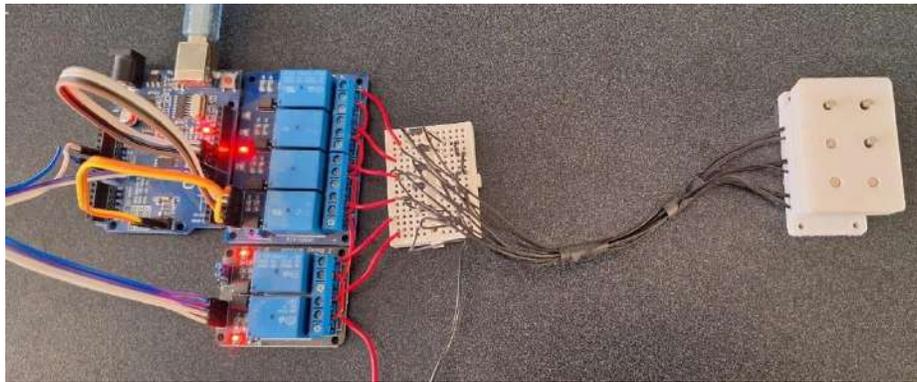


Figura 4-4. Dispositivo de salida representando en la celda braille el signo transmitido por la computadora (Autoría propia, 2024).

4.2 Resultados del sistema

Los resultados que presenta son óptimos y buenos en la gran mayoría de las partes que conforman al sistema. El procesamiento de las imágenes es bueno, el recorte de la imagen para la selección del carácter es correcto, el envío de datos a la tarjeta y representación de los signos braille en la celda electrónica aciertan con el carácter guardado en código ASCII; pero la detección de los caracteres por parte del OCR es donde presenta algunas inconsistencias, es decir, que no logra identificar correctamente a todas las letras, signos de puntuación y números, o en ciertas ocasiones le cuesta detectar ese mismo carácter. Las gráficas que se muestran en páginas subsecuentes son resultados correspondientes al motor OCR, el cual es el Tesseract OCR; esto datos son la eficacia en la identificación de cada letra, signo de puntuación y número de las hojas de pruebas.

❖ Texto en Arial

La gráfica en la Figura 4-5 muestra el porcentaje para la detección de letras minúsculas Arial, en esta imagen se observa que el motor Tesseract OCR solamente logra identificar cinco letras consistentemente, en la letra *h* solo la identifica dos veces y para todas las demás letras, son incorrectamente identificadas; en el caso de las letras *c*, *m*, *n*, *o*, *s*, *u*, *v*, *x*, *y* y *z* las confunde con su contraparte mayúscula. La gráfica mostrada en la Figura 4-6 corresponde con las letras mayúsculas Arial detectadas por el OCR, en esta misma se observa una identificación correcta en gran parte de las letras, solo siete no son identificadas una sola vez, a cinco le cuesta pasar el 50% y otras cuatro están entre el 60% y 80%. Por lo tanto, la detección del motor Tesseract OCR es buena para las letras mayúsculas de esta fuente, no así para la mayoría de las letras minúsculas.



Figura 4-5. Resultados del Tesseract OCR con letras minúsculas en Arial (Autoría propia, 2024).

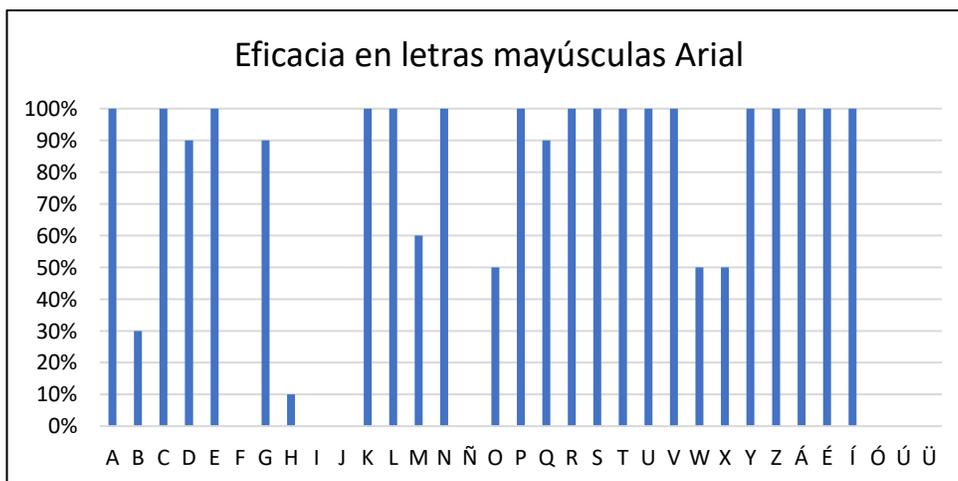


Figura 4-6. Resultados del Tesseract OCR con letras mayúsculas en Arial (Autoría propia, 2024).

En lo que respecta a los números en Arial, el motor Tessract OCR identifica bien a siete de las diez cifras, solo en los números: 2, 7 y 9 no se logra la detección; para el 5 y 8 lo hace hasta un 90%, pero aun así el OCR hace un buen trabajo. En el caso de los signos de puntuación, el OCR solamente detecta al 100% los siguientes signos: », <, >, y los paréntesis de apertura y cierre; con los signos de interrogación obtiene un porcentaje del 40% y 10% respectivamente, pero para todos los demás no logra detectarlos o los confunde con otros caracteres; por lo tanto, el motor Tesseract OCR falla con los signos de puntuación.

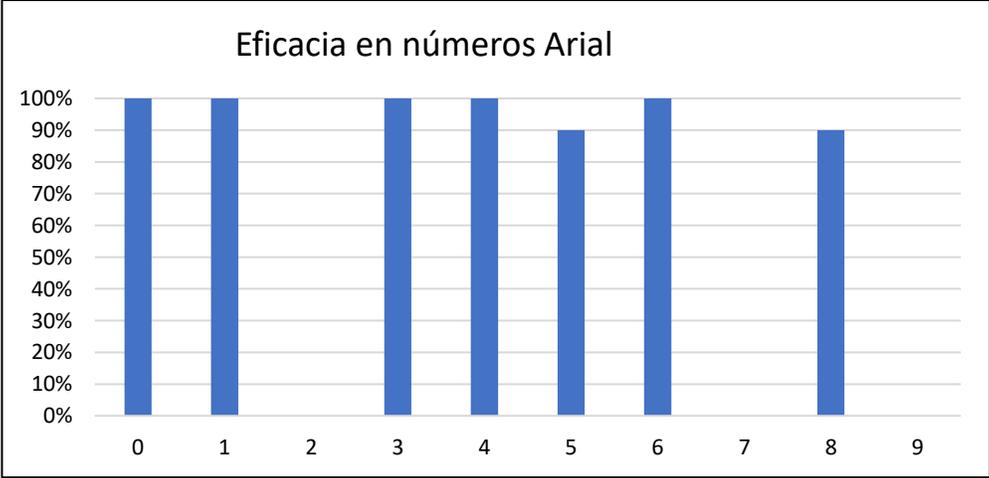


Figura 4-7. Resultado del Tesseract OCR con números Arial (Autoría propia, 2024).

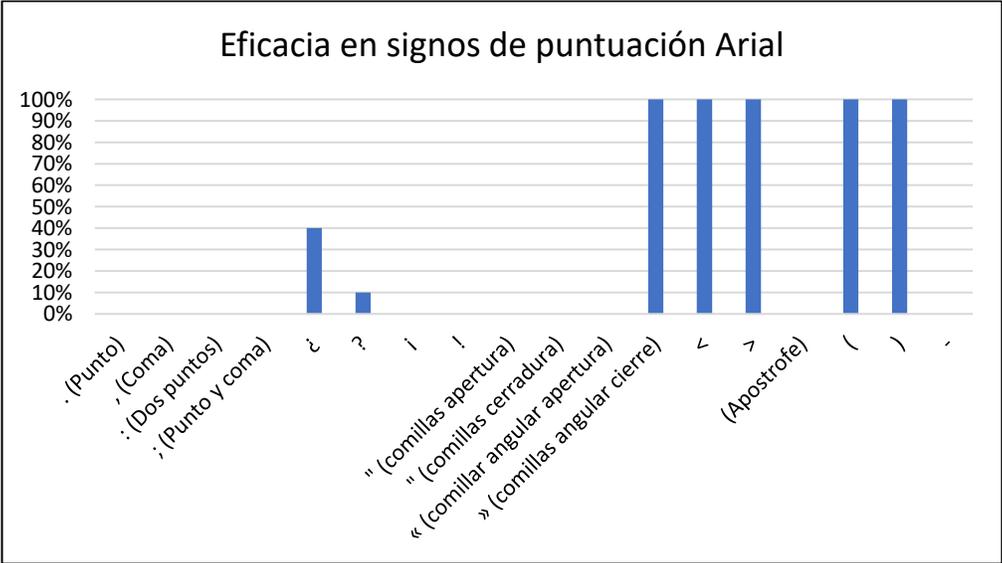


Figura 4-8. Resultados del Tesseract OCR con los signos de puntuación (Autoría propia, 2024).

En lo que respecta a la eficacia con los caracteres con la fuente Arial, esta se puede observar en la Figura 4-9, donde los resultados que se muestran indican que el motor Tesseract OCR tiene un desempeño del 67% con las treinta y tres letras mayúsculas en comparación con el 19% de las letras minúsculas; esto también sucede con los números, con una eficacia del 68%, aunque este valor es solamente para diez caracteres, y en el caso de los signos de puntuación el resultado que se muestra es de 31% para los dieciocho signos. Algo para tener en cuenta en estos resultados es la confusión que el motor OCR tiene para detectar los caracteres, frecuentemente las letras minúsculas que no identifico bien fueron confundidas con letras mayúsculas o en algunos casos con números; esto también sucedió con algunos signos de puntuación, por lo que el motor Tesseract OCR puede estar configurado para detectar de mejor manera las letras mayúsculas y los números con la fuente Arial.

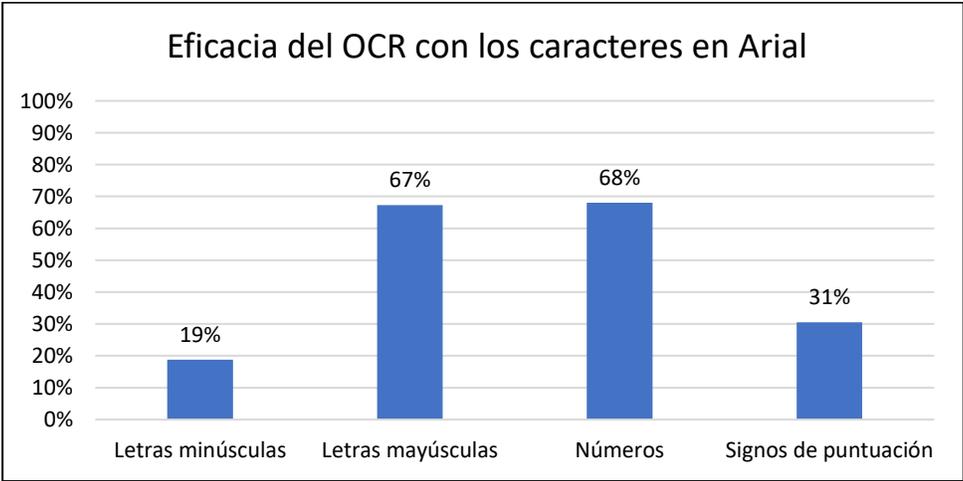


Figura 4-9. Resultados del motor Tesseract OCR con las diferentes categorías de caracteres de la fuente Arial (Autoría propia, 2024).

❖ **Texto en Times New Roman**

Los resultados de la identificación por parte del motor Tesseract OCR en las letras con la fuente Times New Roman, tanto para las minúsculas y como las mayúsculas, se muestran en la Figura 4-10 y Figura 4-11 respectivamente. En las gráficas de las letras minúsculas, se observa que nueve de ellas fueron detectadas al 100%, la é un 80%, v un 40% y unas tres el 10%; el resto no se detectaron, en algunos casos fueron confundidas con sus contrapartes mayúsculas, estas últimas son: *c, m, n, o, s, u, v, w* y *y*. En las gráficas de las letras mayúsculas se observa una mayor cantidad de eficacia, una detección del 100%, en este caso trece, aun así, otras nueve que obtuvieron un resultado de entre 30% y 90%, pero el resto no se detectaron o fueron confundidas con otros caracteres. El OCR tiene problema en detectar la mayoría de las letras minúsculas

de este tipo de fuente. Con las mayúsculas, se obtiene un mejor resultado, aun así, logra fallar.



Figura 4-10. Resultados del Tesseract OCR con letras minúsculas Times New Roman (Autoría propia, 2024).

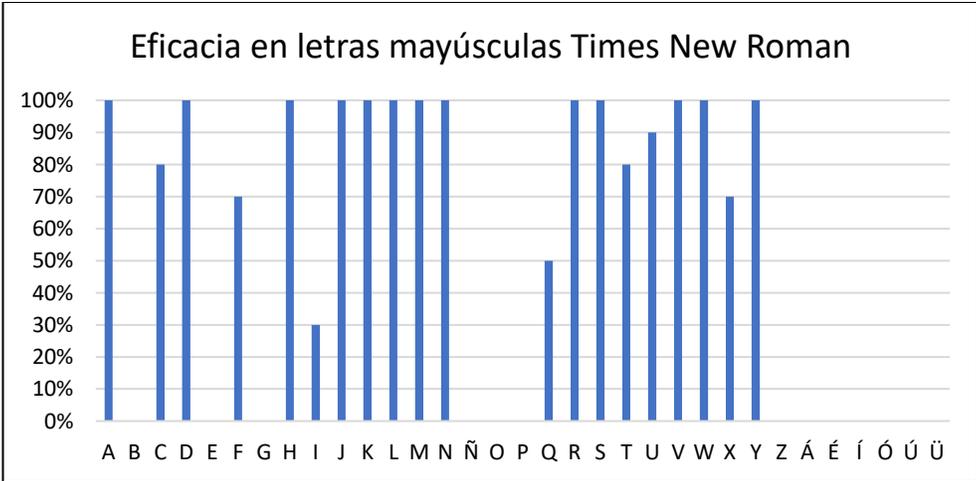


Figura 4-11. Resultados del Tesseract OCR con letras mayúsculas Times New Roman (Autoría propia, 2024).

Los resultados del motor Tesseract OCR con los números, son mostrados en la Figura 4-12, seis de todas las cifras se identificaron correctamente, con un porcentaje del 100%, y las cifras restantes no fueron detectadas ninguna sola vez; por lo que tiene un buen desempeño con este caso, pero aún falla en detectar cuatro de los diez números. El caso de los signos de puntuación, en la gráfica de la Figura 4-13 se puede ver que solo tres caracteres alcanzaron un 100% en la identificación, los cuales son: <, > y el paréntesis de apertura; el paréntesis de cierre tiene un 90%, el signo de interrogación de cierre obtiene un 70%; y el resto, el OCR no logra detectar o los confunde con otros caracteres.

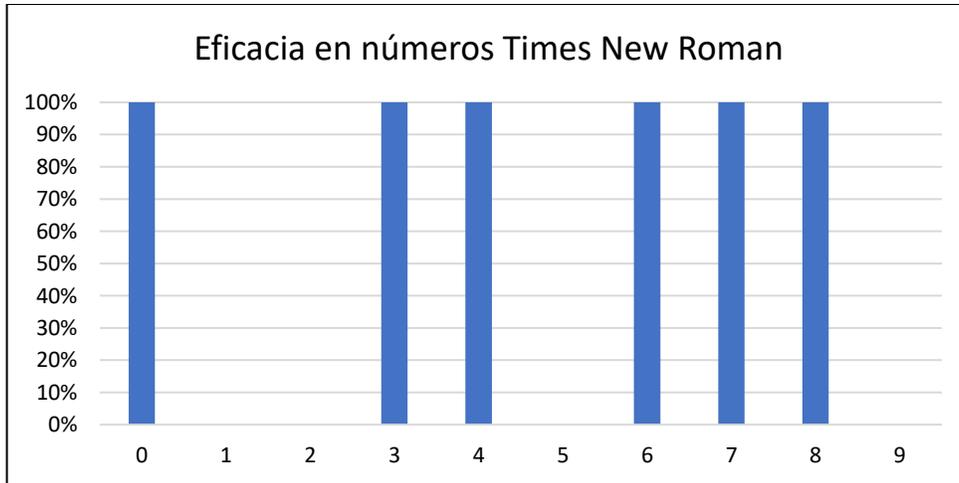


Figura 4-12. Resultados del Tesseract OCR con números Times New Roman (Autoría propia, 2024).



Figura 4-13. Resultados del Tesseract OCR con signos de puntuación Times New Roman (Autoría propia, 2024).

La eficacia del motor Tesseract OCR con las diferentes categorías de los caracteres en Times New Roman se puede ver en la Figura 4-14, en ella los resultados indican que el motor OCR tiene un buen desempeño con los números y las letras mayúsculas, un 60% y 54% de eficacia respectivamente, pero con los signo de puntuación y las minúsculas obtuvo un 31% y 29% respectivamente; esto quiere decir, que el motor OCR realizó un mejor trabajo en estas dos últimas categorías con la fuente Times New Roman que con Arial. Los problemas que presentó el motor OCR para detectar las letras minúsculas con la fuente Times New Roman, fueron los mismos que se manifestaron en Arial, es decir confundir

los caracteres con letras mayúsculas y números, aunque esto no se manifestó con la misma magnitud; por tanto, el funcionamiento del motor OCR con las letras minúsculas fue mejor. En cambio, la eficacia del motor OCR con los signos de puntuación no cambio, ya que se obtuvo el mismo porcentaje, aunque al observar las Figura 4-8 y Figura 4-13, que muestran los resultados para los signos de puntuación en Arial y Times New Roman respectivamente, se puede ver que existe un ligero cambio de porcentaje en algunos caracteres, aun así estos cambios siguen generando el mismo resultado de 31% de eficacia; y en cuanto a la bajo rendimiento se debe a que también continua confundiendo unos signos de puntuación con letra y números; esto último puede deberse a que estos caracteres necesitan de un mayor contexto de letras para ser identificados correctamente, y esto también sucede con los demás caracteres en letras minúsculas, mayúsculas y números.

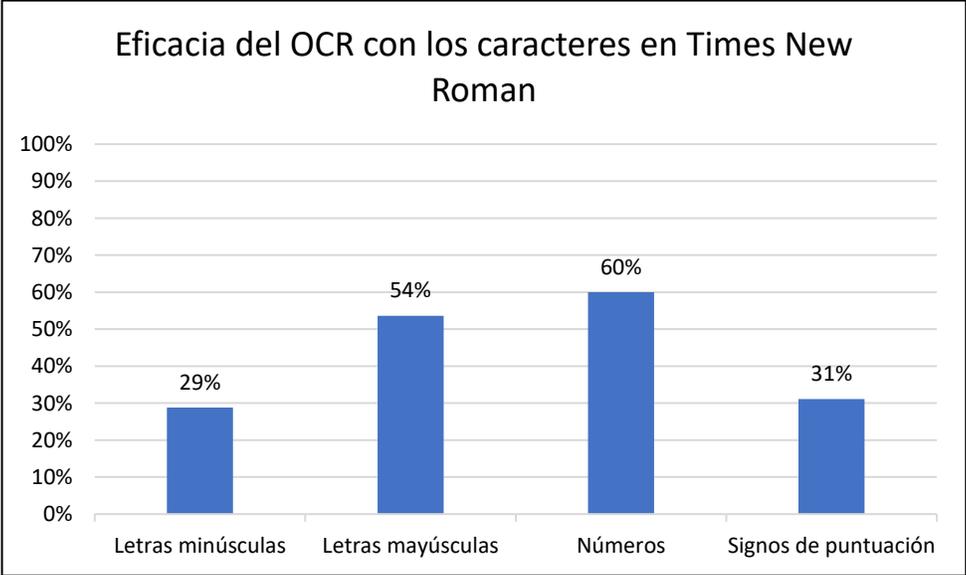


Figura 4-14. Resultados del motor Tesseract OCR con las diferentes categorías de caracteres de la fuente Times New Roman (Autoría propia, 2024).

5 CONCLUSIÓN

El sistema de traducción de texto impreso en español al sistema braille, se logró construir satisfactoriamente con base en los objetivos planteados. Las diferentes partes que componen el sistema funcionan bien de manera individual, esto hablando en la mayoría de las partes; y cuando trabajaron de forma conjunta, lo hicieron de una manera apropiada, es decir, que la comunicación entre las tres partes principales funcionó. Para la primera etapa, que es la adquisición de imágenes, la cámara logró capturar las imágenes con claridad y, se ajustó a la luz del entorno para realizar fotografías con una buena nitidez. La interfaz de usuario, y en sí el programa de computadora, el cual forma la segunda parte del sistema, procesó las imágenes adecuadamente, logrando aplicar distintos métodos para llegar a obtener los resultados deseados en la fotografía, sin embargo, en la parte del reconocimiento de caracteres es donde surgieron detalles, esto se debió a que el motor Tessract OCR y la red neuronal que utiliza no logra identificar el 100% de todos los caracteres en las imágenes; el OCR confundió varias veces unos caracteres con otros, las letras las reconoció como números y viceversa, y con los signos de puntuación no identificó correctamente a la mayoría de estos. En cuanto a la tercera parte, la cual está formada por el dispositivo que representa físicamente a los signos braille, este logró formar los signos braille de los caracteres que el OCR identificó, incluso si el carácter no correspondía con lo mostrado en la imagen, es decir, el OCR entregaba una letra, un número o un signo de puntuación, después le seguía la conversión de este a un signo braille y su posterior representación en la celda electrónica, la cual era correcta, esto sucedía independientemente de si el carácter era el que la cámara capturó o no, por lo tanto esta última etapa del sistema funcionó correctamente.

En términos de los objetivos planteados en este trabajo, estos se cumplieron. La interfaz de comunicación entre una computadora y una cámara digital se desarrolló correctamente, esto utilizando el lenguaje de programación C#, la cámara envía las imágenes y la computadora las recibe para ser mostradas y procesadas en el programa de computadora. La identificación de los caracteres en las imágenes, por medio de un el motor Tesseract OCR con redes neuronales se completó, aunque este no logró detectar correctamente todas las letras, números y signos que se encontraban en las imágenes teniendo mejor desempeño en las letras Arial mayúsculas. La transcripción de los caracteres a sistema braille se concluyó correctamente, esto se debió a que se siguieron las normas de la CBE para representación de texto español en signos braille. El objetivo de programar una tarjeta y, que esta controle un circuito construido se terminó satisfactoriamente; esto permitió controlar los solenoides, los cuales forman los puntos de los signos braille en una celda electrónica actualizable. El diseño y construcción de una celda actualizable braille se logró crear, esto con la ayuda de las herramientas CAD y de la fabricación por medio de manufactura aditiva. Las pruebas de lectura y escritura, para el sistema de traducción de texto impreso a braille, se realizaron adecuadamente, corroborando el funcionamiento de este mismo.

La construcción de las tres partes principales planteadas en la metodología, y los componentes de los que se conforman, permiten entender de manera más profunda el funcionamiento de cada una de ellas. En la primera etapa, que es la de adquisición y donde se encuentra la cámara digital, se necesitó investigar la manera en cómo trabaja este dispositivo, las limitaciones que posee y el proponer una solución a los problemas que se presentaron; este fue el caso con la iluminación del entorno y, cómo, ésta afecta a la captura de imágenes, por lo que se necesitó implementar un control de auto iluminación en la cámara. También estuvo el hecho de que las capturas debían estar a una distancia fija, ya que el enfoque del aparato es manual, por lo que se terminó diseñando y fabricando una base con medidas fijas y óptimas para la cámara.

La creación de la segunda parte del sistema, la cual es el programa que procesa las imágenes, está también planteó varios retos para ser creada. Entre los problemas que se tuvieron en esta etapa, uno fue el entender la forma de usar correctamente los diferentes métodos de procesamiento, y la secuencia que estos debían seguir para obtener imágenes útiles en le reconociendo por parte del OCR; pero el otro problema más desafiante, fue el de utilizar las características que las imágenes entregaron, es decir, el plantear una serie de algoritmos que utilizaran correctamente las características de los objetos en las imágenes, esto con el propósito de corregir la imagen, o el de seleccionar las partes que forman un letra, signo o número. La programación para la conversión de los caracteres en código ASCII a signo braille, se realizó usando las normas de la CBE para cada letra, signo y número como una base de datos, esto permitió que la transformación de los caracteres fuera automática; una vez con la información del signo braille, éste simplemente es enviado a la tarjeta con el microcontrolador.

El diseño y construcción de la celda braille presentó el problema en las dimensiones de los solenoides, esto provocó que se rediseñara varias veces la base donde se colocan los actuadores, por ende, también se tuvo que fabricar unas cuantas bases, hasta que los solenoides encajaron adecuadamente. El tema del control para los solenoides fue cubierto por la tarjeta con el microcontrolador, en esta se programó un algoritmo capaz de interpretar de una forma correcta la información, la cual era enviada desde la computadora por el programa que procesa las imágenes. El control de los solenoides se planteó para representar dos tipos de signos braille: uno que solamente muestra signos simples, que son aquellos que son un signo braille; y los segundos, los cuales están compuesto por dos signos braille, como lo son los números, letras mayúsculas y unos cuantos signos de puntuación. Al final de la construcción de la tercera etapa del sistema, se presentó un problema, el cual fue la alimentación del circuito de potencia y la energía que requieren los solenoides para ser activado; ya que estos últimos necesitan de una cierta corriente, la cual, termina provocando un calentamiento en estos mismos si su periodo de operación es prolongado, pero aun así funcionan sin mayor inconveniente durante las pruebas del sistema.

5.1 Prospectivas del trabajo

Los datos obtenidos en las pruebas del sistema de traducción de texto a braille, sobre todo en la parte del reconocimiento de caracteres, permiten plantear mejoras en el sistema. Primero, se tiene el mejorar el proceso de identificación por parte del motor Tesseract OCR; después realizar pruebas con los ajustes, ajustarlo hasta que reconozca correctamente el 100% de las letras, signos y números de manera individual, con esto mencionado, llevar el sistema de traducción a reconocer en imágenes, texto, el cual debe estar formado por palabras y caracteres de manera conjunta, e incluso en diferentes renglones; esto último requerirá ajustar el programa que mejora las imágenes y el algoritmo que selecciona el carácter a representar en la celda braille electrónica.

La celda braille electrónica implementada en este trabajo, requiere usar otro tipo de actuadores, ya que los solenoides se calientan por estar operando durante un largo tiempo; con unos nuevos actuadores que no consuman una gran energía, pero sobre todo corriente, se lograría resolver este problema. Los nuevos actuadores podrían tener otro tipo de dimensiones, por lo que se debe diseñar una nueva base para estos mismos, y si es posible ajustar el tamaño de esta nueva base con los tamaños recomendados por la CBE, para que la nueva celda braille electrónica represente los signos braille con las dimensiones correctas, es decir, que el tamaño y distribución de los puntos de esta nueva celda braille corresponda con los de un signo braille normal, el cual está definido por las normas de la CBE.

En otro aspecto a futuro, de este trabajo, es llevar este desarrollo a un sistema más pequeño y portátil, el cual pueda ser utilizado desde una sola mano y que no requiera un computador grande. Las mejoras que este nuevo sistema para la traducción texto a braille, están: una nueva tarjeta con microcontrolador, la cual sea capaz de procesar imágenes de manera óptima; reducir el tamaño de la celda braille electrónica, esto con el propósito de que el sistema se ajuste a la mano del usuario, y que la celda pueda utilizarse en los dedos. Un último punto por mejorar es unir la cámara digital con la celda braille electrónica, esto para tener en una sola base dos partes del sistema, la cual permitirá tener el dispositivo de adquisición y de salida en un solo punto; además esta nueva base podrá ajustarse a la mano del usuario, de tal forma que le permitirá a éste, usar el dispositivo en la lectura de texto impreso, pero de una manera como si estuviera leyendo en braille.

6 BIBLIOGRAFÍA

- Acevedo, M. E., Alonso, J., Martínez, F., & Neri, K. A. (2016). Associative models applied to a Braille-Spanish/Spanish-Braille translator. *2016 8th Euro American Conference on Telematics and Information Systems (EATIS)*. <https://doi.org/10.1109/EATIS.2016.7520138>
- Adnan, M. E., Dastagir, N. M., Jabin, J., Chowdhury, A. M., & Islam, M. R. (2017). A Cost Effective Electronic Braille for Visually Impaired Individuals. *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, 175–178. <https://doi.org/10.1109/R10-HTC.2017.8288932>
- Aguado Behar, A. (2000). *Temas de identificación y control adaptable*. Instituto de Cibernética, Matemática y Física.
- Başçiftçi, F., & Eldem, A. (2016). An interactive and multi-functional refreshable Braille device for the visually impaired. *Displays*, *41*, 33–41. <https://doi.org/10.1016/j.displa.2015.11.001>
- Borovikov, E. (2004). *A survey of modern optical character recognition techniques*. [Archivo PDF]. <https://doi.org/10.13140/2.1.2690.2404>
- Devi, V. A. (2016). Portable Device Converting Braille and Text Document to Speech Output Using Raspberry Pi. *International Journal of Control Theory and Applications (IJCTA)*, 4809–4818. https://www.researchgate.net/publication/309067093_Portable_device_converting_braille_and_text_document_to_speech_output_using_raspberry_pi
- García Soria, F., & Ruiz Prieto, P. (2014). *Documento técnico B 1: Parámetros dimensionales del braille* (Versión 1). ONCE. <https://www.once.es/servicios-sociales/braille/comision-braille-espanola/documentos-tecnicos/documentos-tecnicos-relacionados-con-el-braille/documentos-tecnicos-relacionados-con-el-braille>
- García Soria, F., & Ruiz Prieto, P. (2023). *Documento técnico B 2: Signografía básica de las lenguas cooficiales españolas* (F. J. Martínez Calvo (ed.); Versión 3). ONCE. <https://www.once.es/servicios-sociales/braille/comision-braille-espanola/documentos-tecnicos/documentos-tecnicos-relacionados-con-el-braille/documentos-tecnicos-relacionados-con-el-braille>
- Gibson, I., Rosen, D., Stucker, B., & Khorasani, M. (2021). *Additive Manufacturing Technologies* (3th Ed.). Springer.
- Gomis-Bellmunt, O., & Campanile, L. F. (2010). *Design Rules for Actuators in Active Mechanical Systems*. Springer.
- González, C. C. (2016). *Diseño y construcción de una plantilla para sistema Braille*. Universidad Nacional Autónoma de México.
- Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing*. Pearson.
- Griffing, L. R. (2023). *Imaging Life: Image Acquisition and Analysis in Biology and Medicine*. John Wiley and Sons.

- Hassan, K. M. N., Biswas, S. K., Anwar, M. S., Siam, M. S. I., & Shahnaz, C. (2019). A Dual-Purpose Refreshable Braille Display Based on Real Time Object Detection and Optical Character Recognition. *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*, 78–81. <https://doi.org/10.1109/SPICSCON48833.2019.9065110>
- Hossain, S., Raied, A. A., Rahman, A., Abdullah, Z. R., Adhikary, D., Khan, A. R., Bhattacharjee, A., Shahnaz, C., & Fattah, S. A. (2018). Text to Braille Scanner with Ultra Low Cost Refreshable Braille Display. *2018 IEEE Global Humanitarian Technology Conference (GHTC)*. <https://doi.org/10.1109/GHTC.2018.8601552>
- Hussain, M. A., Rifat, R. I., Iqbal, S. B., Biswas, S., Alam, M. G. R., & Reza, M. T. (2022). Deep Learning based Bangla Voice to Braille Character Conversion System. *2022 IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. <https://doi.org/10.1109/IEMCON56893.2022.9946619>
- IEEE. (1990). *IEEE Standard Glossary of Software Engineering Terminology*. IEEE. <https://ieeexplore.ieee.org/document/159342>
- INEGI. (2020). *Población con discapacidad o limitación en la actividad cotidiana por entidad federativa y tipo de actividad realiza según sexo, 2020*. Discapacidad. https://www.inegi.org.mx/app/tabulados/interactivos/?pxq=Discapacidad_Discapacidad_02_2c111b6a-6152-40ce-bd39-6fab2c4908e3&idrt=151&opc=t
- Jothi, G., Azar, A. T., Qureshi, B., & Kamal, N. A. (2022). IReader: An Intelligent Reader System for the Visually Impaired. *2022 7th International Conference on Data Science and Machine Learning Applications (CDMA)*, 188–193. <https://doi.org/10.1109/CDMA54072.2022.00036>
- Karmel, A., Sharma, A., Pandya, M., & Garg, D. (2019). IoT based Assistive Device for Deaf, Dumb and Blind People. *Procedia Computer Science*, 165, 259–269. <https://doi.org/10.1016/j.procs.2020.01.080>
- Kobayashi, A., Onishi, J., Nishizaki, H., & Kitaoka, N. (2022). End-to-End Speech to Braille Translation in Japanese. *2022 IEEE International Conference on Consumer Electronics (ICCE)*. <https://doi.org/10.1109/ICCE53296.2022.9730468>
- Kumari, S., Akole, A., Angnani, P., Bhamare, Y., & Naikwadi, Z. (2020). Enhanced Braille Display Use of OCR and Solenoid to Improve Text to Braille Conversion. *2020 International Conference for Emerging Technology (INCET)*, 1–5. <https://doi.org/10.1109/INCET49848.2020.9153996>
- Li, Z., Wang, R., Zhang, T., Xu, X., & Liang, P. (2019). Intelligent Braille Conversion System of Chinese Characters Based on Markov Model. *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 1283–1287.

<https://doi.org/10.1109/ITNEC.2019.8729267>

- Liu, Z., Luo, Y., Cordero, J., Zhao, N., & Shen, Y. (2016). Finger-Eye: A Wearable Text Reading Assistive System for the Blind and Visually Impaired. *2016 IEEE International Conference on Real-Time Computing and Robotics (RCAR)*, 123–128. <https://doi.org/10.1109/RCAR.2016.7784012>
- Lo, S.-W., & Chou, H.-M. (2022). Evaluating and Improving Optical Character Recognition (OCR) Efficiency in Recognizing Mandarin Phrases with Phonetic Symbols. *2022 IEEE International Conference on Internet of Things and Intelligence Systems (IoTIS)*, 390–394. <https://doi.org/10.1109/IoTIS56727.2022.9975969>
- Martínez Calvo, F. J. (2006). *Características de la rotulación para personas con discapacidad visual* (F. J. Martínez Calvo (ed.)). Organización Nacional de Ciegos Españoles (ONCE).
- Navarro Saad, M. (1998). *El sistema braille*. Eureka [Archivo PDF]. <https://www.uaq.mx/ingenieria/publicaciones/eure-uaq/n13.html>
- Pressman, R. S. (2010). *Ingeniería del Software, un enfoque práctico* (7th Ed.). Mc Graw Hill.
- Rahimi, M., Shen, Y., Peng, C., Liu, Z., & Jiang, F. (2022). Opto-electrotactile Feedback Enabled Text-line Tracking Control for A Finger-wearable Reading Aid for the Blind. *2022 International Conference on Robotics and Automation (ICRA)*, 7007–7013. <https://doi.org/10.1109/ICRA46639.2022.9811610>
- Ramos-García, O. I., Vuelvas-Alvarado, A. A., Osorio-Pérez, N. A., Ruiz-Torres, M. Á., Estrada-González, F., Gaytan-Lugo, L. S., Fajardo-Flores, S. B., & Santana-Mancilla, P. C. (2022). An IoT Braille Display towards Assisting Visually Impaired Students in Mexico †. *Engineering Proceedings*, 27(1), 1–5. <https://doi.org/10.3390/ecsa-9-13194>
- Rao, S. S. (2009). *Engineering Optimization Theory and Practice* (4th Ed.). John Wiley and Sons.
- Simón, C., Ochaíta, E., & Huertas, J. A. (1995). El sistema Braille: Bases para su enseñanza- aprendizaje. *Comunicación, Lenguaje y Educación*, 91–102. <https://doi.org/10.1174/021470395763771891>
- Sultana, S., Rahman, A., Chowdhury, F. H., & Zaman, H. U. (2017). A Novel Braille Pad with Dual Text-to-Braille and Braille-to-Text capabilities with an integrated LCD Display. *2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, 195–200. <https://doi.org/10.1109/ICICICT1.2017.8342559>
- Toyserkani, E., Sarker, D., Ibhádode, O. O., Liravi, F., Russo, P., & Taherkhani, K. (2022). *Metal Additive Manufacturing*. John Wiley and Sons.
- Valdéz Pérez, F. E., & Pallás Areny, R. (2007). *Microcontroladores: Fundamentos y Aplicaciones con PIC*. Alfaomega.

- Zaman, S., Abrar, M. A., Hassan, M. M., & Islam, A. N. M. N. (2019). A Recurrent Neural Network Approach to Image Captioning in Braille for Blind-Deaf People. *2019 IEEE International Conference on Signal Processing, Information, Communication and Systems, SPICSCON 2019*, 49–53. <https://doi.org/10.1109/SPICSCON48833.2019.9065144>
- Zamir, M. F., Khan, K. B., Khan, S. A., & Rehman, E. (2020). Smart Reader for Visually Impaired People Based on Optical Character Recognition. *Communications in Computer and Information Science*, 79–89. https://doi.org/10.1007/978-981-15-5232-8_8

7 APÉNDICES

7.1 Apéndice A: Artículo del dispositivo de salida

Dispositivo Electromecánico de Puntos en Relieve Actualizable para Reproducción de Caracteres del Sistema Braille

Ing. Jorge Agustín Velázquez Villegas¹, Dr. Miguel Trejo Hernández², Dr. Roque Alfredo Osornio Rios³, Dr. Irving Armando Cruz Albarrán⁴, Dr. Martin Valtierra Rodríguez⁵ y Dr. J. Jesús de Santiago Pérez⁶.

Resumen— El acceso a la información es algo imprescindible para que una persona pueda desarrollarse. El medio por el cual se transmite también es un aspecto importante, por eso, el sistema Braille de lectoescritura es una manera en que las personas invidentes son capaces de leer. En este trabajo se presenta el desarrollo de un sistema que permite traducir texto digital en español a caracteres en sistema Braille. Este proceso se realiza por medio de la intercomunicación entre una interfaz de computadora, un microprocesador y un dispositivo electromecánico conformado por actuadores, dispuestos en una celda Braille e implementada con manufactura aditiva. Este dispositivo permite que se actualicen los puntos en relieve del sistema Braille de acuerdo con el carácter a representar, es así como un usuario puede verificar el carácter formado en la celda. El desempeño del sistema es bueno para traducir palabras en párrafos escritos desde la interfaz desarrollada.

Palabras clave— Braille, traducción, celda braille, español.

Introducción El sistema braille español

El signo Braille

En la Figura 7-1, de acuerdo con Calvo (2006), se muestra la forma de representar un signo Braille, en ella los puntos en negro indican cuales deben resaltar sobre una superficie plana para ser percibidos, en este caso los seis puntos estarán en alto. Este formato tiene 64 combinaciones para simbolizar las distintas letras del alfabeto español, los signos de puntuación, signos auxiliares y, dependiendo la forma en que estén dispuestas una celda después de otra, la secuencia representará un número, una letra en mayúscula u otros signos especiales. Los puntos de un signo Braille tienen una numeración, siendo los de la columna izquierda del 1 al 3 en orden descendente, y los de la columna derechas del 4 al 6, igualmente descendente.

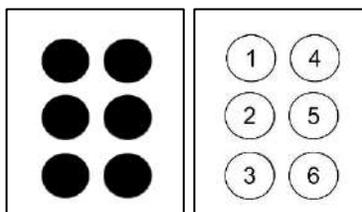


Figura 7-1. Signo Braille de seis puntos y su numeración (Calvo, 2006).

¹ Ing. Jorge Agustín Velázquez Villegas es estudiante de la Maestría en Ciencias Mecatrónica en la Universidad Autónoma de Querétaro. jvelazquez31@alumnos.uaq.mx

² Dr. Miguel Trejo Hernández es Profesor-Investigador de la Universidad Autónoma de Querétaro. mtrejo@hspdigital.org (autor correspondiente).

³ Dr. Roque Alfredo Osornio Rios es Profesor-Investigador de la Universidad Autónoma de Querétaro. raosornio@hspdigital.org

⁴ Dr. Irving Armando Cruz Albarrán es Profesor-Investigador de la Universidad Autónoma de Querétaro. irving.cruz@uaq.mx

⁵ Dr. Martin Valtierra Rodríguez es Profesor-Investigador de la Universidad Autónoma de Querétaro. martin.valtierra@uaq.mx

⁶ Dr. J. Jesús de Santiago Pérez es Profesor-Investigador de la Universidad Autónoma de Querétaro. sapji@uaq.mx

Signos Braille en español

Los signos braille son definidos por la CBE (Comisión Braille Española) (2023), para tener un estándar en el idioma español, las 27 letras del alfabeto tienen su conversión al sistema Braille, incluso las vocales con acento y la *ü* se representan con su propio signo braille, esto se observa en la Figura 7-2a, los números debajo de cada símbolo indican los puntos de la celda que deben formar un relieve, como se vio en la Figura 7-1. Los signos de puntuación también poseen un signo para ser identificados en el sistema Braille, la imagen de la Figura 7-2b muestra su conversión y el número de los que sobresalen en la superficie, una observación a estos es que el signo de interrogación de apertura y cerradura son el mismo en Braille, igualmente es el caso para los signos de admiración, las comillas dobles y las comillas simples.

a	b	c	d	e	f	g	h	i	j
•	••	•••	••••	•••••	••••••	•••••••	••••••••	•••••••••	••••••••••
1	12	14	145	15	124	1245	125	24	245
k	l	m	n	ñ	o	p	q	r	s
••	•••	••••	•••••	••••••	•••••••	••••••••	•••••••••	••••••••••	•••••••••••
13	123	134	1345	12456	135	1234	12345	1235	234
t	u	v	w	x	y	z			
•••	••••	•••••	••••••	•••••••	••••••••	•••••••••			
2345	136	1236	2456	1346	13456	1356			
á	é	í	ó	ú		ü			
••••	•••••	••••••	•••••••	••••••••		•••••••••			
12356	2346	34	346	23456		1256			

(a)

••••••••••	3	•
••••••••••	2	•
••••••••••	23	;
••••••••••	25	:
••••••••••	3-3-3	...
••••••••••	26	¿?
••••••••••	235	¡!
••••••••••	236	“» ««
••••••••••	6-236	“<>”
••••••••••	126	(
••••••••••	345)

(b)

Figura 7-2. Signos Braille para el idioma español: (a) signos Braille de las letras minúsculas del alfabeto; (b) signos Braille de los signos de puntuación (CBE, 2023).

Los números en Braille requieren del signo que representa un número, este es mostrado en la Figura 7-3a, cada dígito del 1 al 9 utiliza los signos de la *a* hasta la *i* respectivamente para simbolizar tal número y, el 0 implementa el signo de la *j*; escribir el ‘1’ en Braille requiere primero utilizar el signo del número y después el signo de la letra *a*, otro ejemplo es 203, donde primero se escribe el signo del número y para seguir con los otros tres signos correspondientes al dígito.

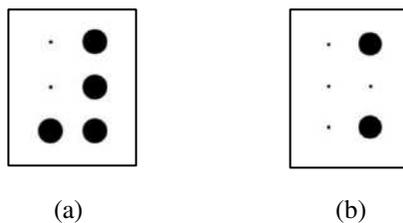


Figura 7-3. Signos Braille: (a) signo del número; (b) signo para la letra mayúscula (CBE, 2023).

Las letras mayúsculas en Braille son un caso similar a los números, requieren del signo mayúscula que se observa en la imagen (b) de la Figura 7-3, sin embargo, este signo puede ser utilizado de dos formas: la primera es colocarlo antes de cada signo para cada letra, esta opción genera que una palabra en mayúsculas tenga el doble de signos Braille; la segunda forma es escribir al principio de una palabra en mayúsculas el signo mayúscula, para posteriormente seguir con los signos de las letras.

Metodología

Proceso de traducción

El diagrama metodológico empleado en el trabajo se muestra en la Figura 7-4. El primer paso de la metodología consiste en escribir texto en la interfaz de software, este programa, desarrollado en Visual Studio, identifica los caracteres y los convierte a los signos Braille correspondientes con cada letra, signo de puntuación o número, basándose en las codificaciones ASCII de cada carácter y una base de conocimiento en el mismo. El siguiente paso es la conexión entre la computadora y la tarjeta con el microcontrolador basado en Arduino Uno, los dos dispositivos reciben y envían datos el uno al otro. Subsecuente a lo anterior el microcontrolador interpreta la información de entrada y la transforma en señales, estas últimas se mantendrán por tiempo determinado, el cual es elegido desde el programa de computadora, al terminarse este periodo, la tarjeta mandará un dato a la computadora para que la aplicación envíe otro signo Braille, o bien, termine el proceso. El penúltimo paso consiste en activar los relevadores de un circuito de potencia con las señales del microcontrolador, esto permite que los actuadores electromecánicos sean alimentados con el suficiente voltaje para ser accionados; por último, los solenoides se mueven de arriba para abajo dependiendo de las señales y se mantienen, en la posición, hasta que el tiempo marcado por el programa de computadora termine. El tiempo que puede elegir el usuario varía entre 1 a 5 segundos dependiendo de la habilidad y conocimiento para la lectura de los caracteres por parte de él.

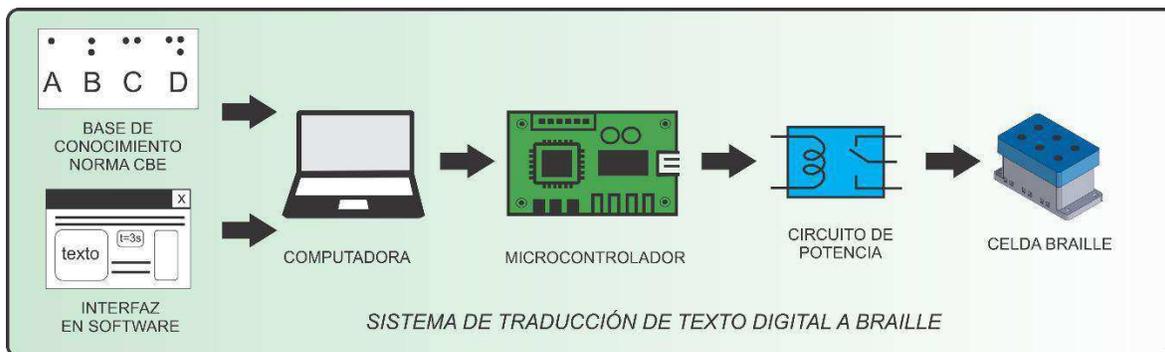


Figura 7-4. Metodología del sistema de traducción.

Sistema de traducción

El sistema que se ha desarrollado convierte un texto digital en español a Braille de manera continua, el texto es representado en una celda con un arreglo de actuadores electromecánicos, estos mismos, son accionados de acuerdo con el signo Braille que simboliza una letra o carácter del texto. El sistema se compone de dos partes, la primera consiste en una interfaz de software desarrollada en Visual Studio, este se muestra en la imagen (a) de la Figura 7-5, y la segunda parte es un microcontrolador (Arduino Uno) para accionar a través de una etapa de potencia compuesta por relevadores de 5V, los 6 solenoides, a su vez, son de DC12V, 0,2 A, con potencia de 2,4W y un peso de 11g cada uno. Estos últimos dispuestos en un arreglo de 6 en un elemento impreso con manufactura aditiva, se observa en la imagen (b) de la misma Figura 7-5.

En la interfaz, el usuario escribe el texto en el cuadro de edición, o bien, lo pega de otra fuente. Este es convertido a Braille cuando el botón de “Traducir texto” es presionado; Se anexa también un menú de opciones de tiempo “Tiempo de espera para el signo”, es el periodo en el que los actuadores se mantendrán en una posición, las opciones a elegir van de 1 a 5 segundos. Debajo del menú se muestra el renglón donde se encuentra el carácter que es enviado a la tarjeta, la palabra que lo contiene y el carácter de texto en sí. La transmisión del texto convertido es continua, por lo tanto, el usuario solamente presiona el botón de traducción en el software una sola vez.

El microcontrolador recibe un dato que representa un signo Braille, la tarjeta lo convierte en una cadena de 6 bits y, dependiendo del signo Braille, el valor de cada bit será 0 o 1; estos valores se extraen de la cadena y son convertidos a señales eléctricas, mismas que activarán los relevadores de la etapa de potencia

correspondientes y al mismo tiempo permiten alimentar a cada actuador para que cambien de posición, elevando o bajando su vástago, cuya carrera es de 4mm.

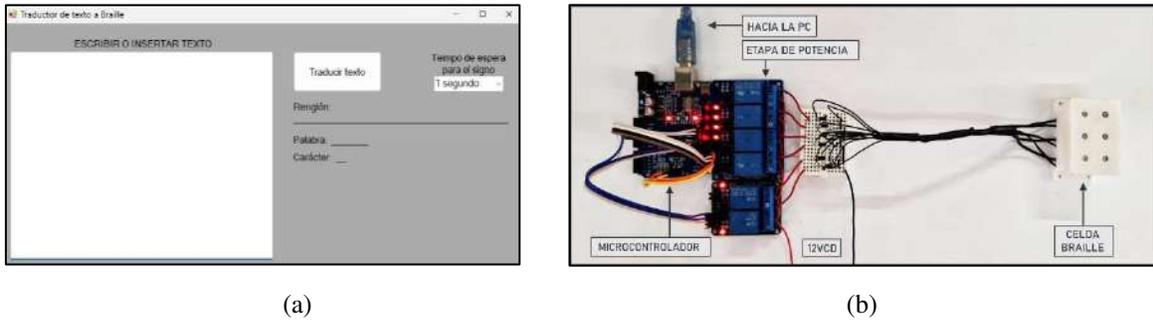


Figura 7-5. Sistema: (a) Interfaz de software; (b) Dispositivos electrónicos y electromecánicos para la celda Braille.

Resultados

La repuesta del sistema tuvo buenos resultados de acuerdo con cada valor de tiempo elegido en la interfaz. El accionamiento de la celda respondió de manera adecuada acorde a los caracteres que se transmitieron a la tarjeta. El tiempo de representación de cada signo estuvo acorde a la opción seleccionada en el menú, mismo que debe oscilar entre 2 y 3 segundos si se tiene conocimiento adecuado de los símbolos, o de 4 a 5 segundos si apenas se comienza con el aprendizaje del método. Como ejemplo, uno de los textos probados en el sistema se muestra en la Figura 7-6, las palabras y los signos que se tienen como evidencia se encuentran en la Figura 7-7. En ésta misma, se muestran las palabras, su representación en Braille, y la evidencia fotográfica de cada signo empleado para representar las palabras que aparecen en el escrito: *Sueño, rocío con una coma al final y alegre*. Como se puede observar, cada letra tiene su correspondiente signo Braille en la celda, así como también, el signo para una letra mayúscula se reprodujo al principio de la palabra *Sueño*.

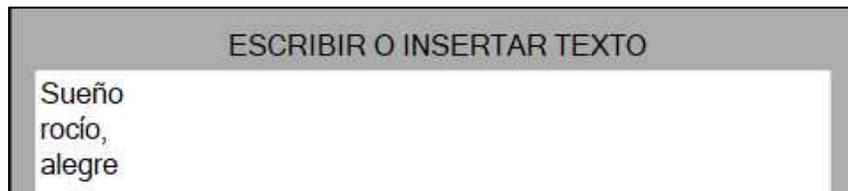


Figura 7-6. Texto de prueba para la celda Braille.

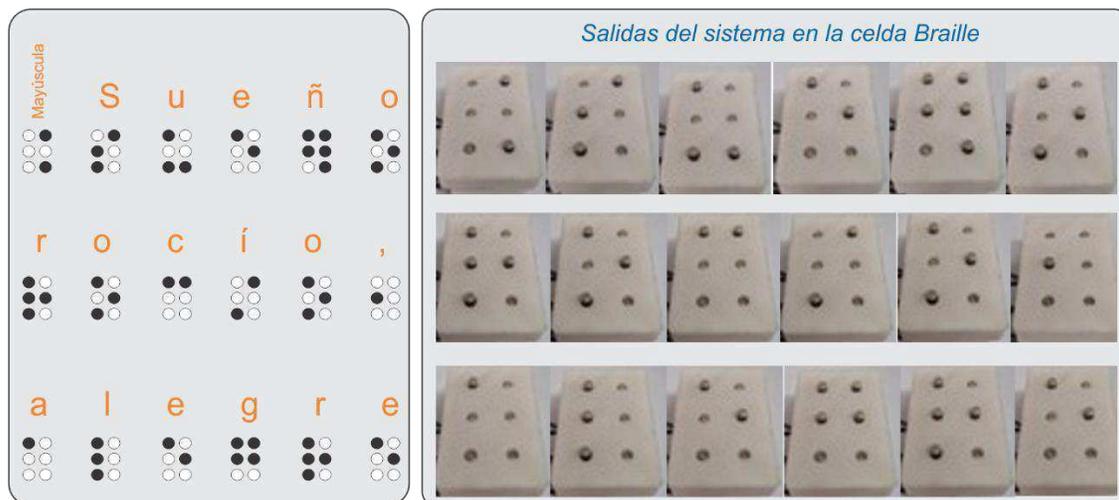


Figura 7-7. Resultados de la traducción de caracteres de las palabras en la celda Braille.

Discusión

Como se pudo observar, el sistema de traducción de caracteres, a partir de texto digital, muestra buen desempeño en la traducción cuando se envía texto en diversos renglones. El manejo de los tiempos brinda flexibilidad al usuario dependiendo del nivel de aprendizaje o conocimiento del sistema. Durante el entrenamiento en el conocimiento de los signos, el usuario elige la rapidez en la muestra de los signos en la celda Braille desarrollada y la puede ir aumentando de acuerdo con su habilidad. El consumo de corriente por parte del sistema hace que éste se caliente si el tiempo de uso es prolongado, por lo que un área de oportunidad es diseñar una nueva carcasa, y producirla en manufactura aditiva, que incluya un sistema de enfriamiento con un pequeño ventilador. La miniaturización de la celda es otro reto importante, sin embargo, el buen desempeño, hasta ahora, en las pruebas realizadas es muy notoria.

Conclusión

El aporte del presente trabajo se enfoca principalmente en la traducción de caracteres de texto vía software-hardware para su reproducción en un dispositivo electromecánico, con una matriz de puntos de relieve de 3x2, actualizable en la altura de los puntos en relieve, de acuerdo con el carácter a representar, como una manera de apoyo en el aprendizaje y lectura de caracteres en español en sistema Braille. A partir del uso y traducción de una base de conocimiento de los signos Braille, los resultados en la utilización del sistema, muestran un excelente desempeño, aunado a que se puede indicar el tiempo en el que los signos se reproducen dependiendo de la habilidad del usuario. No toda la información ni conocimiento se encuentra disponible en formato Braille, debido a que el acceso a este tipo de materiales de lectura-escritura es escaso, o los que existen suelen tener un alto costo. De las personas con alguna discapacidad en México, alrededor del 44% cuentan con alguna discapacidad visual según datos del INEGI (Instituto Nacional de Estadística y Geografía) (2020), y a la mayoría, está tecnología es aún desconocida, por lo que es de suma importancia no dejar de investigar y producir conocimiento en estos ámbitos.

Referencias

- Adnan, M. E., Dastagir, N. M., Jabin, J., Chowdhury, A. M., y Islam, M. R. “A cost effective electronic braille for visually impaired individuals”. *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, IEEE , 175–178. Dhaka, Bangladesh. 2017.
- Calvo, F. J. “Características de la rotulación para personas con discapacidad visual”. *Organización Nacional de Ciegos Españoles (ONCE)*, Madrid, España. 2006.
- Comisión Braille Española (CBE), “Documento técnico B 2: Signografía básica de las lenguas cooficiales españolas”. *Documentos técnicos relacionados con el Braille* (en línea), Versión 3, 2023, consultada por Internet el 16 de agosto de 2023. Dirección de internet: <https://www.once.es/servicios-sociales/braille/comision-braille-espanola/documentos-tecnicos/documentos-tecnicos-relacionados-con-el-braille/documentos-tecnicos-relacionados-con-el-braille>
- Hassan, K. M. N., Biswas, S. K., Anwar, M. S., Siam, M. S. I., & Shahnaz, C. “A Dual-Purpose Refreshable Braille Display Based on Real Time Object Detection and Optical Character Recognition”, *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*, IEEE, 78–81, Dhaka, Bangladesh. 2019.
- Hossain, S., Raied, A. A., Rahman, A., Abdullah, Z. R., Adhikary, D., Khan, A. R., Bhattacharjee, A., Shahnaz, C., & Fattah, S. A. “Text to Braille Scanner with Ultra Low Cost Refreshable Braille Display”, *2018 IEEE Global Humanitarian Technology Conference (GHTC)*, IEEE, 1–6, San Jose, CA, EE. UU., 2018.
- Instituto Nacional de Estadística y Geografía (INEGI). “Población con discapacidad o limitación en la actividad cotidiana por entidad federativa y tipo de actividad realiza según sexo, 2020” *Discapacidad* (en línea), 2020, consultada por Internet el 22 de febrero de 2023. Dirección de internet: https://www.inegi.org.mx/app/tabulados/interactivos/?pxq=Discapacidad_Discapacidad_02_2c111b6a-6152-40ce-bd39-6fab2c4908e3&idrt=151&opc=t

Kumari, S., Akole, A., Pallavi, A., Yash, B., & Zaid, N. “Enhanced Braille Display Use of OCR and Solenoid to Improve Text to Braille Conversion”, *2020 International Conference for Emerging Technology (INCET)*, IEEE, 1–5, Belgaum, India, 2020.

Liu, Z., Luo, Y., Cordero, J., Zhao, N., & Shen, Y. “Finger-eye: A wearable text reading assistive system for the blind and visually impaired”, *2016 IEEE International Conference on Real-Time Computing and Robotics (RCAR)*, IEEE, 123–128, Angkor Wat, Camboya, 2016.

Rahimi, M., Shen, Y., Peng, C., Liu, Z., & Jiang, F. “Opto-electrotactile Feedback Enabled Text-line Tracking Control for A Finger-wearable Reading Aid for the Blind”, *2022 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 7007–701, Filadelfia, PA, EE. UU., 2022.

Ramos-García, O. I., Vuelvas-Alvarado, A. A., Osorio-Pérez, N. A., Ruiz-Torres, M., Estrada-González, F., Gaytan-Lugo, L. S., Fajardo-Flores, S. B., & Santana-Mancilla, P. C. “An IoT Braille Display towards Assisting Visually Impaired Students in Mexico”, *Engineering Proceedings*, MDPI, Volumen 27 ECSA-9, 1–5, 2022.

Sameia, Z., Abid, A. M., Muntasir, H. M., & Islam, A. N. M. N. “A Recurrent Neural Network Approach to Image Captioning in Braille for Blind-Deaf People”, *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*, IEEE, 49–53, Dhaka, Bangladesh, 2019.



CONGRESO INTERNACIONAL DE INVESTIGACIÓN ACADEMIA JOURNALS HIDALGO 2023

Arte, Tecnología, Equidad

CERTIFICADO

otorgado a

Ing. Jorge Agustín Velázquez Villegas
Dr. Miguel Trejo Hernández
Dr. Roque Alfredo Osornio Ríos
Dr. Irving Armando Cruz
Dr. Martín Valtierra Rodríguez
Dr. J. Jesús de Santiago Pérez

por su artículo titulado

Dispositivo Electromecánico de Puntos en Relieve Actualizable para Reproducción de Caracteres del Sistema Braille

(Artículo No. HLG261)

La ponencia correspondiente fue presentada en el Congreso Internacional de Investigación Academia Journals Hidalgo 2023, *Arte, Tecnología, Equidad*, desarrollado los días 18, 19, y 20 de octubre de 2023, co-organizado por el PYSEIP, *Productos y Servicios Estratégicos para Investigación y Posgrado*, y la Universidad Politécnica Metropolitana de Hidalgo. El artículo está incluido en las siguientes publicaciones:

- (1) en el portal de Internet AcademiaJournals.com, con ISSN 1946-5351 online, Vol. 15, No. 09, 2023 e indexación en la base de datos *Fuente Académica Plus de EBSCOHOST*, Massachusetts, Estados Unidos y
- (2) en libros ebook digitales compilados por área temática, con números ISBN online*. Se tiene acceso libre a todas las publicaciones del congreso en el portal de internet de Academia Journals.

Los organizadores del congreso reconocen la participación de los autores en el congreso, agradeciendo sus contribuciones.

DR. RAFAEL MORAS, P.E.

Director General
Academia Journals



7.2 Apéndice B: Texto de pruebas en Arial

a b c d e f g h i j k l m

n ñ o p q r s t u v w x y

z á é í ó ú ü

A B C D E F G H I J K L M

N Ñ O P Q R S T U V W X Y

Z Á É Í Ó Ú Ü

. , : ; ¿ ? ¡ “ ” « » <

> ’ () -

0 1 2 3 4 5 6 7 8 9

7.3 Apéndice C: Texto de pruebas en Times New Roman

a b c d e f g h i j k l m
n ñ o p q r s t u v w x y
z á é í ó ú ü
A B C D E F G H I J K L M
N Ñ O P Q R S T U V W X Y
Z Á É Í Ó Ú Ü
. , : ; ¿ ? ¡ “ ” « » <
> ’ () -
0 1 2 3 4 5 6 7 8 9