

UNIVERSIDAD SANTO TOMÁS

Software de respuesta empática para robots

Realizado por

Juan David Avendaño Garzón

Proyecto de Grado presentado en cumplimiento del requisito
para optar por el grado de Ingeniería Electrónica



Grupo de Investigación GED (Grupo de Estudio y Desarrollo en Robótica)
Facultad de Ingeniería Electrónica
División de Ingenierías

Junio de 2025

Software de respuesta empática para robots

Realizado por

Juan David Avendaño Garzón

Proyecto de Grado presentado en cumplimiento del requisito
para optar por el grado de Ingeniería Electrónica

Dirigido por

Ing. Armando Mateus Rojas M.Sc

Grupo de Investigación GED (Grupo de Estudio y Desarrollo en Robótica)
Facultad de Ingeniería Electrónica
División de Ingenierías

Junio de 2025

Autoridades de la Universidad

RECTOR GENERAL

R.P. FRAY ALVARO JOSÉ ARANGO RESTREPO, O.P.

VICERRECTOR ADMINISTRATIVO Y FINANCIERO GENERAL

R.P. FRAY HENDER ALVEIRO RODRIGUEZ PÉREZ, O.P.

VICERRECTOR ACADÉMICO GENERAL

R.P. FRAY MAURICIO ANTONIO CORTES GALLEGO, O.P.

SECRETARIO GENERAL

Dra. INGRID LORENA CAMPOS VARGAS

DECANO DIVISIÓN DE INGENIERÍAS

R.P. FRAY JAVIER ANTONIO HINCAPIÉ ARDILA, O.P.

SECRETARIA DE DIVISIÓN

E.C. LUZ PATRICIA ROCHA CAICEDO

DECANO FACULTAD DE INGENIERÍA ELECTRÓNICA

ING. LEONARDO FABIO YEPES ARBELÁEZ, PhD.

Nota de aceptación

Firma del autor

Firma del jurado

Firma del jurado

BOGOTÁ D.C. — DE 2025

Advertencia

La Universidad Santo Tomás no se hace responsable de las opiniones y conceptos expresados en el trabajo de grado, solo velará por que no se publique nada contrario al dogma ni a la moral católica y porque el trabajo no tenga ataques personales y únicamente se vea el anhelo de buscar la verdad científica.

Capítulo III –Art. 46 del Reglamento de la Universidad Santo Tomás.

Dedicatoria

Para mi yo de 13 años. Este es el primer paso de nuestro sueño más grande. Ha valido la pena todo lo que hemos vivido hasta hoy y vamos por más.

Agradecimientos

Quiero agradecer profundamente mi mejor amiga Alejandra Pinzón quien me ha apoyado con todos sus conocimientos y siempre creyendo en mí para poder desarrollar este proyecto.

A mi director Armando Mateus Rojas quien ha creído en mí y me ha apoyado en la realización de este proyecto.

A mi papá, mi hermana, mi abuelita, mi tía y mi primo quienes a pesar de la adversidad siempre estuvieron conmigo en todo momento apoyándome y colaborándome en lo que más pudieran.

A mis amigos y compañeros Juan, J. Larrotta, Sergio y D. Santiago quienes me apoyaron incondicionalmente, estuvieron en mis buenos y malos ratos, siempre motivándome a ser mejor y aportando con sus conocimientos para poder lograr culminar con este proyecto.

Al psicólogo Néstor Mejía y a la futura psicóloga Dania León quienes me colaboraron no solo a entender la parte psicológica del proyecto, sino también a fortalecer y afianzar mi determinación a realizar este proyecto.

A mis profesores y anteriores jurados quienes aportaron con sus críticas constructivas y sus conocimientos para el desarrollo y mejora de este proyecto.

Me gustaría extender un agradecimiento a mi profesor del colegio Rolando Barajas por haberme instruido y llevado a esta carrera.

Por último, pero no menos importante, gracias Dios por todas estas personas que me diste para poder cumplir con este sueño que hoy en día es una realidad.

Resumen

Este proyecto de grado tiene como objetivo desarrollar un software que permita a un robot social (como el caso del robot Pepper con el que cuenta la Universidad Santo Tomás) responder de forma empática ante diferentes emociones humanas, enfocándose principalmente en la alegría y la tristeza. La idea nace de una necesidad clara: los robots actuales aún no logran establecer una conexión emocional real con las personas, lo que limita mucho su papel en entornos sociales.

Para lograr una interacción más natural y cercana, se diseñó un software a base modelo conceptual de empatía que le permite al robot identificar emociones a través de tres canales principales: reconocimiento facial, reconocimiento de voz y detección de palabras clave. A cada uno se le asignó un peso específico dentro de un sistema de votación ponderada, priorizando el canal que ofrece mayor confiabilidad según estudios psicológicos y pruebas internas.

El desarrollo se llevó a cabo utilizando ROS, conectando un nodo externo con Pepper para garantizar una comunicación efectiva sin depender completamente del sistema del robot. Durante las pruebas, realizadas en entornos controlados, el sistema logró reconocer emociones con buenos niveles de precisión, especialmente en la emoción de alegría. Sin embargo, también se evidenciaron algunas dificultades, como la sensibilidad a la iluminación y al ruido, o el uso de accesorios como gafas, que pueden afectar el reconocimiento facial.

A pesar de estas limitaciones, se comprobó que el sistema funciona y cumple con el objetivo de ofrecer una interacción más empática. Este proyecto no solo fortalece la línea de investigación del Grupo GED en robótica social, sino que también abre la posibilidad de aplicar este tipo de soluciones en el acompañamiento emocional o la salud mental (bajo la supervisión y orientación de profesionales idóneos).

Abstract

This undergraduate project aims to develop software that enables a social robot (such as the Pepper robot available at Santo Tomás University) to respond empathetically to different human emotions, focusing primarily on joy and sadness. The idea stems from a clear need: current robots still fail to establish a genuine emotional connection with people, which greatly limits their role in social environments.

To achieve a more natural and engaging interaction, the software was designed based on a conceptual model of empathy, allowing the robot to identify emotions through three main channels: facial recognition, voice recognition, and keyword detection. Each channel was assigned a specific weight within a weighted voting system, prioritizing the most reliable one according to psychological studies and internal testing.

Development was carried out using ROS, connecting an external node to Pepper to ensure effective communication without fully relying on the robot's system. During controlled testing, the system successfully recognized emotions with good accuracy, particularly in detecting joy. However, some challenges were also observed, such as sensitivity to lighting, noise, or accessories like glasses, which can affect facial recognition.

Despite these limitations, the system was proven functional and met the goal of delivering a more empathetic interaction. This project not only strengthens the research line of the GED Group in social robotics but also opens the possibility of applying such solutions in emotional support or mental health (under the supervision and guidance of qualified professionals).

Índice general

List of Figures	XI
List of Tables	XII
Glosario	XIII
1. Introducción	1
1.1. Planteamiento del problema	1
1.2. Justificación	2
1.3. Objetivos	3
1.3.1. Objetivo general	3
1.3.2. Objetivos específicos	3
1.4. Impacto social	4
2. Estado del arte	5
2.1. IA para la empatía	5
2.2. Psicología de la empatía en IA	6
2.3. Estudios y proyectos sobre reconocimiento de emociones	7
2.4. Proyectos base	10
3. Diseño metodológico	12
3.1. Fase 1: Investigación	12
3.2. Fase 2: Selección de herramientas de reconocimiento de emociones	13
3.3. Fase 3: Construcción del software	13
3.4. Fase 4: Prueba	14
3.5. Fase 5: Documentación	14
4. Marco teórico	15
4.1. Aspectos de teoría psicológica	15
4.1.1. Modelo conceptual de empatía	15
4.1.2. Respuesta empática	16
4.2. Herramientas de reconocimiento de emociones	17
4.2.1. Herramientas de reconocimiento facial	17
4.2.2. Herramienta de reconocimiento de voz	18
4.3. Robótica	19

4.3.1.	ROS	19
4.3.2.	Lenguaje de programación	20
4.3.3.	Plataforma robótica: Pepper	20
5.	Desarrollo conceptual	21
5.1.	Investigación	21
5.1.1.	Modelo conceptual de empatía establecido	21
5.1.1.1.	Respuesta de validación	22
5.1.2.	Reconocimiento de emociones	24
5.1.3.	Herramientas de reconocimiento	24
5.1.4.	ROS	25
5.2.	Selección de herramientas de reconocimiento de emociones	27
5.2.1.	Reconocimiento facial	28
5.2.2.	Reconocimiento de voz	30
5.2.3.	Usos en la plataforma robótica Pepper	32
5.3.	Construcción del software	33
5.3.1.	Diagrama de flujo	33
5.3.2.	Nodo de ROS	35
5.3.3.	Implementación de herramientas	35
5.3.3.1.	Reconocimiento facial (face emotion recognition):	35
5.3.3.2.	Reconocimiento de voz y palabras (voice and word recognition):	35
5.3.4.	Selección de respuesta según emoción	36
5.4.	Pruebas	38
5.4.1.	Conexión con pepper por ROS	38
5.4.2.	Reconocimiento facial	38
5.4.3.	Reconocimiento de voz	39
5.4.4.	Votación ponderada	41
5.4.5.	Software completo	42
5.5.	Documentación	44
5.5.1.	Repositorio	44
6.	Resultados y discusión	45
6.1.	Resultados	45
6.1.1.	Establecer un modelo conceptual de empatía aplicable a la interacción humano - robot, para generar respuestas afectivas y cognitivas basadas en procesos existentes de clasificación de emociones.	45
6.1.2.	Implementar un software basado en ROS para la interacción humano – robot mediante la integración del modelo conceptual de empatía y el sistema de clasificación de emociones.	46
6.1.2.1.	Reconocimiento de la empatía (Empathy Recognition)	46
6.1.2.2.	Proceso (Process)	50
6.1.2.3.	Resultados (Outcome)	51
6.1.2.4.	Respuesta empática (Empathic Response)	51

6.1.3. Validar el software mediante el uso de casos de prueba y métricas de desempeño en entornos controlados para evidenciar el correcto funcionamiento usando la plataforma robótica Pepper.	52
6.1.3.1. Ideales	53
6.1.3.2. Reales	55
6.1.3.3. Comparación de datos	57
6.1.4. Discusión	57
7. Conclusiones	59
Bibliografía	61
Anexos de Implementación	69
.1. Conexión y control del robot Pepper	69
.2. Reconocimiento de emociones por voz con Pepper	71

Índice de figuras

1.	Modelo conceptual de empatía para HRI	16
2.	Maya facial	18
3.	Ilustración voz	19
4.	Lógica de la Herramienta 1	28
5.	Demostración de la herramienta 2	29
6.	Evidencia del reconocimiento facial: Alegría	33
7.	Evidencia del reconocimiento facial: Tristeza	33
8.	Flujo del sistema de validación emocional	34
9.	Conexión con pepper por medio de ROS establecida	38
10.	Reconocimiento facial a imagen	39
11.	Evidencia de correcto funcionamiento de reconocimiento de voz: Resultado	40
12.	Evidencia error 1: Reconocimiento de voz con chatbots	40
13.	Evidencia error 2: Reconocimiento de voz con chatbots	40
14.	Caso prueba 1: Entradas iguales	41
15.	Caso prueba 2: Entradas distintas	41
16.	Caso prueba 3: 2 Entradas con menor peso mismo valor	41
17.	Caso prueba 4: 2 Entradas con mayor peso mismo valor	42
18.	Evidencia final: Conexión, reconocimiento de palabras y de voz	43
19.	Evidencia final: Reconocimiento facial a una imagen	43
20.	Matriz de confusión normalizada	46
21.	Matriz de correlación de pares	47
22.	Conjunto de gráficos combinados: bounding boxes y Mapas de dispersión	48
23.	Curva de F1 frente a la confianza	48
24.	Uso de la herramienta Roboflow	50
25.	Diagrama de flujo del sistema	73

Índice de cuadros

1.	Comparación técnica entre Pepper y un computador con Ubuntu 20.04 respecto a compatibilidad con ROS Noetic	26
2.	Evaluación de herramientas de reconocimiento facial según los requerimientos	30
3.	Comparación técnica de parámetros acústicos entre emociones vocales: alegría vs. tristeza	32
4.	Pesos asignados a cada uno de los reconocimientos de emociones	36
5.	Resultados del análisis de emoción (evidencia: EvidenciaRecoTristezaExitosa.jpg)	39
6.	Resultados del análisis de emoción detectada	44
7.	Limitaciones de Pepper	52
8.	Resultados del análisis de emoción detectada en niño con expresión triste	53
9.	Resultados del análisis de emoción detectada en hombre feliz con tono de piel oscuro	54
10.	Resultados del análisis de emoción detectada en joven con expresión de sorpresa	54
11.	Resultados del análisis multimodal de emoción en interacción con la palabra “feliz”	55
12.	Resultados del análisis de emoción detectada en imagen con dos personas	55
13.	Resultados del análisis de emoción detectada en persona con gafas	56
14.	Resultados del análisis de emoción detectada en condiciones de baja luz	56
15.	Resultados del análisis multimodal de emoción en interacción con la palabra “feliz”	57

Glosario

- **AAM (Active Appearance Model)** Modelo de apariencia activa utilizado para la detección y análisis de características faciales en el reconocimiento de emociones.
- **ADAS (Advanced driver-assistance systems)** Sistemas avanzados de asistencia al conductor, que permiten mejorar no solo la seguridad del vehículo que lo posee, sino también la de los demás.
- **API (Application Programming Interface)** Interfaz de programación de aplicaciones que permite la interacción entre diferentes programas o componentes de software.
- **ANNs (Artificial Neural Networks)** Redes neuronales artificiales utilizadas para el reconocimiento y clasificación de patrones, aplicadas comúnmente en la detección de emociones y aprendizaje profundo.
- **DBT (Dialectical Behavior Therapy)** Terapia Dialéctico-Conductual utilizada en psicología, enfocada en la validación emocional y el bienestar del paciente.
- **HMM (Hidden Markov Models)** Modelos de Markov Ocultos utilizados para el modelamiento estadístico de secuencias, especialmente en reconocimiento de voz y emociones.
- **HRI (Human-Robot Interaction)** Interacción Humano-Robot, campo de estudio que analiza la comunicación y cooperación entre humanos y robots.
- **IA (Inteligencia Artificial)** Rama de la informática que busca desarrollar sistemas capaces de realizar tareas que requieren inteligencia humana.
- **IAPS (International Affective Picture System)** Sistema internacional de imágenes afectivas utilizado en estudios psicológicos para inducir y categorizar emociones.
- **NN (Neural Network)** Red neuronal, modelo computacional inspirado en el cerebro humano, utilizado en aprendizaje automático.
- **ODS3 (Objetivo de Desarrollo Sostenible 3)** Objetivo de la ONU enfocado en asegurar la salud y el bienestar en todas las etapas de la vida.

- **OpenCV (Open Source Computer Vision Library)** Biblioteca de visión artificial de código abierto utilizada para el procesamiento de imágenes y videos.
- **OS (Operating System)** Sistema operativo que gestiona los recursos de hardware y software de un computador.
- **ROS (Robot Operating System)** Sistema operativo para robots que facilita el desarrollo de aplicaciones robóticas mediante nodos y mensajes.
- **SVM (Support Vector Machine)** Algoritmo de aprendizaje supervisado usado para clasificación, útil en reconocimiento de emociones faciales.
- **UA (Unidades de Acción)** Movimientos musculares faciales específicos que permiten analizar expresiones emocionales, según la teoría de Paul Ekman.

Capítulo 1

Introducción

1.1. Planteamiento del problema

Durante el siglo XXI, los avances tecnológicos han transformado diversos campos, impulsando el desarrollo humano a través de innovaciones sistemáticas. Entre estos progresos destacan el desarrollo de sensores avanzados, la percepción multimodal, la visión por computadora y el aprendizaje por refuerzo. Particularmente, la IA y la robótica han registrado un crecimiento notable en años recientes. Sin embargo, a pesar de estos logros, los robots diseñados para interactuar en entornos sociales aún enfrentan una limitación significativa, como es la falta de comunicación genuinamente empática con las personas, la cual trata de la capacidad de interpretar, comprender y responder adecuadamente las emociones y sentimientos [1].

La empatía se define como la capacidad de entender y comprender los sentimientos y emociones con los demás [2], siendo un elemento fundamental en las interacciones sociales humanas. De esta forma, el entendimiento implica captar el significado de algo, mientras que la comprensión va más allá, ya que requiere interiorizar y apropiarse de ese conocimiento, lo que conduce a una comprensión más profunda. No obstante, esta capacidad de comprensión profunda es algo que las máquinas aún no han logrado experimentar, debido a que carecen de la vivencia emocional necesaria para interpretar y asimilar plenamente las emociones humanas [3].

Es por esto, que una de las principales dificultades para desarrollar respuestas empáticas en robots radica en la diversidad de definiciones del término "empatía" como se evidencia en las investigaciones [4], [5] y [6]. Esta falta de consenso ha complicado el establecimiento de un criterio unificado sobre lo que constituye una respuesta empática en un robot. Por consiguiente, la definición que se adoptará en este proyecto es la propuesta por [6], que describe la empatía

como "una respuesta afectiva que surge de la aprehensión o comprensión de los sentimientos, el estado o la condición emocional del otro, y que llega a ser similar a lo que la otra persona está sintiendo o podría llegar a sentir". En otras palabras, se trata de una reacción emocional basada en la comprensión de los sentimientos de los demás.

A pesar de los avances en la interacción humano - robot (HRI, por sus siglas en inglés), aún no se ha logrado una comunicación empática que sea completamente natural y adaptable a cualquier situación. Un ejemplo de esta limitación se observa en las respuestas inadecuadas del robot Pepper de SoftBank Robotics: ante la afirmación "no estoy bien", el robot respondió "me alegra que los dos estemos bien"; esta situación pone en manifiesto la falta de respuestas empáticas y resalta la necesidad de desarrollar estrategias más sofisticadas y efectivas para integrar la empatía en los robots sociales. En este contexto, surge la siguiente pregunta problema:

¿Cómo implementar una respuesta empática en robots sociales para generar interacciones humano - robot que generen respuestas afectivas y cognitivas positivas en los humanos?

1.2. Justificación

A pesar de los avances significativos en HRI implementado en la robótica social, como la capacidad de mantener conversaciones autónomas, aún persiste una importante brecha en el desarrollo de respuestas empáticas genuinas y naturales basadas en la detección de emociones. Esto se debe, a que los robots carecen de una comprensión real de las emociones humanas.

Por ejemplo, investigaciones realizadas en la Universidad Santo Tomás han explorado sistemas HRI que utilizan lenguaje multimodal [7]. Estos sistemas han demostrado mejoras en las respuestas en tiempo real, ya que ofrecen respuestas más acordes al contexto de la interacción. Además, se han logrado avances en la integración de diversas formas de comunicación, como el lenguaje verbal y los gestos. Sin embargo, estos estudios aún no abordan a profundidad el establecimiento de relaciones empáticas.

En este contexto, el presente trabajo se centra en la recreación de respuestas empáticas en un robot, empleando el reconocimiento de emociones y la generación de respuestas adecuadas. Según la definición de empatía adoptada para este proyecto, comprender las emociones de otro individuo [8] implica experimentarlas de manera similar. No obstante, dado que un robot no puede experimentar emociones de forma genuina, se ha optado por simular comportamientos empáticos mediante el procesamiento del lenguaje natural. Para lograrlo, se ha combinado el

análisis de palabras y textos (similar al de un *chatbot* [9]) con el reconocimiento e interpretación de expresiones faciales [10].

Una vez que se identifica el estado emocional de la persona, el robot genera una respuesta empática hacia el usuario. Por ello, se desarrolló un protocolo situacional predefinido que permite al robot responder de manera natural ante determinadas emociones, simulando empatía y promoviendo una HRI más cómoda. Este protocolo [11] se basa en una serie de respuestas predeterminadas según la emoción detectada, diseñadas con fundamentos psicológicos como la escucha activa, con el fin de garantizar una conversación más fluida y empática [12].

Además, este proyecto no solo responde al interés de mejorar la HRI, sino que también contribuye al Grupo de Estudio y Desarrollo en Robótica (GED) al explorar un área poco desarrollada en sus robots sociales, permitiendo avanzar en el desarrollo de plataformas robóticas más adaptados a interacciones reales con personas, un desafío clave en el campo [13] [14].

1.3. Objetivos

1.3.1. Objetivo general

Desarrollar un software basado en ROS para la interacción humano – robot empleando un modelo conceptual de empatía.

1.3.2. Objetivos específicos

- Establecer un modelo conceptual de empatía aplicable a la interacción humano - robot, para generar respuestas afectivas y cognitivas basadas en procesos existentes de clasificación de emociones.
- Implementar un software basado en ROS para la interacción humano – robot mediante la integración del modelo conceptual de empatía y el sistema de clasificación de emociones.
- Validar el software mediante el uso de casos de prueba y métricas de desempeño en entornos controlados para evidenciar el correcto funcionamiento usando la plataforma robótica Pepper.

1.4. Impacto social

Este proyecto contribuye directamente al Objetivo de Desarrollo Sostenible 3 (ODS3), “Salud y bienestar”, cuyo propósito es asegurar una vida sana y promover el bienestar en todas las etapas de la vida. En particular, se alinea con la meta 3.4, enfocada en reducir la mortalidad prematura por enfermedades no transmisibles y en fortalecer la salud mental. En este contexto, el software desarrollado busca brindar apoyo emocional a personas que experimentan tristeza, facilitando la expresión de sus emociones a través de una interacción empática con un robot social [15].

Más allá del beneficio individual, esta iniciativa aporta a la construcción de una sociedad más resiliente frente a los desafíos emocionales cotidianos. Al incorporar una HRI que establece una comunicación más natural y coherente, lo que permite generar vínculos significativos entre las personas y la tecnología. Esto abre nuevas posibilidades para acompañar procesos emocionales en tiempo real y de manera no invasiva, utilizando herramientas innovadoras para abordar el cuidado de la salud mental [16].

Así mismo, este proyecto se articula con lo establecido en el documento “Marco de Proyección Social y Extensión Universitaria” de la Universidad Santo Tomas mediante la estrategia de “Desarrollo comunitario” la cual busca que los programas, sin importar la modalidad, diseñen y generen proyectos en un marco real de aplicación del conocimiento para la comunidad, dando como resultado soluciones alternativas a las problemáticas que cada una de estas tiene [17].

Desde la universidad, este tipo de proyectos encuentran sentido al surgir del diálogo entre el conocimiento académico y las realidades sociales, culturales y económicas de las comunidades. Así, el desarrollo tecnológico no solo responde a una necesidad técnica, sino que se convierte en una herramienta concreta para el cambio, alineada con el compromiso ético y social de la institución.

Capítulo 2

Estado del arte

La integración de la empatía en la robótica social se ha vuelto esencial para mejorar la HRI, especialmente en áreas como la educación y el cuidado de personas mayores. Esto ha impulsado la investigación sobre cómo representar, detectar y simular la empatía en sistemas basados en IA o en plataformas robóticas.

Se realizó una revisión de artículos y proyectos base usando herramientas como Google Scholar e IEEE Xplore, lo que permitió identificar documentos clave que abordan tanto la definición como los retos de implementar la empatía en IA y en la robótica.

El estado del arte se organiza en tres áreas principales: la implementación técnica de IA para detectar emociones, la contribución de la psicología en el diseño de sistemas empáticos, y los estudios sobre el reconocimiento de emociones en la HRI.

2.1. IA para la empatía

Desde la perspectiva de la Ingeniería Electrónica, fue fundamental investigar proyectos previos que pudieran contribuir de manera significativa a la implementación de este trabajo. Para ello, se seleccionaron artículos que abarcaban desde el reconocimiento facial hasta el uso de *Deep Learning*. Entre estos, se destacó un artículo que describía el desarrollo de un sistema de reconocimiento de emociones a través del análisis facial mediante IA, utilizando algoritmos como Eigenfaces y librerías como Numpy, OpenCV y Sklearn [18]. Este trabajo presentó, en primer lugar, cómo identificar emociones a partir de rasgos faciales.

Además, se exploró cómo las plataformas de aprendizaje adaptativo podían aprovechar el reconocimiento facial para ofrecer experiencias educativas más personalizadas. La integración de esta tecnología permitió que la IA identificara expresiones que indicaban falta de comprensión por parte del alumno, proporcionando un apoyo adicional y motivador [19]. Esta personalización no solo mejoró la experiencia de aprendizaje, sino que también contribuyó a reducir la deserción estudiantil en dichas plataformas, evidenciando una posible aplicación del proyecto en el ámbito educativo.

Los avances en la detección de emociones fueron fundamentales para comprender y fomentar el desarrollo de habilidades empáticas, especialmente en la integración de robots en roles sociales [20]. Esto evidenció cómo la tecnología puede simular la empatía y mejorar la HRI.

Por otro lado, se investigaron aplicaciones relacionadas con el análisis de rasgos psicológicos y cognitivos a través de la escritura y el dibujo a mano. En este contexto, se discutió cómo los modelos de aprendizaje profundo contribuyeron a la detección temprana de enfermedades cognitivas mediante la evaluación de dibujos manuales. Aunque estos desarrollos aún se encontraban en fase de evaluación, sugirieron que la IA podía ofrecer nuevas herramientas para detectar signos tempranos de dichas enfermedades, lo que resultaría en una mejora en la atención médica y la calidad de vida de las personas [21].

Estos avances en el reconocimiento de emociones y rasgos psicológicos evidencian el progreso de la tecnología en su capacidad para comprender y responder a las necesidades humanas. Esto proporciona bases sólidas para el desarrollo del proyecto y demuestra que su implementación representa más ventajas que desventajas.

2.2. Psicología de la empatía en IA

Una de las bases de este proyecto es la psicología. Al ser una ciencia que surge del deseo humano por comprender la mente y la conducta humana, nos permite entender el concepto “empatía” y cómo esta se podría desarrollar en un robot. El progreso continuo en la investigación psicológica ha impulsado la aplicación de estos conocimientos en la creación de sistemas inteligentes, fomentando el desarrollo de campos como la psicología cognitiva. En el artículo [22] se observa una investigación más detallada que nos lleva a entender la relación entre la psicología y la IA, enfatizando en que esta última debe ser considerada como una herramienta complementaria y no un sustituto de la intervención humana.

Además, existe una estrecha relación entre la neurociencia y la psicología, siendo ambas fundamentales para los avances en computación y el nacimiento de la IA. Sin embargo, persiste la incertidumbre sobre cómo se integrará la IA y la psicología en el futuro [23], lo que resalta otra rama de la psicología que se involucra con las tecnologías emergentes, permitiendo evidenciar su uso como una herramienta.

Asimismo, se observan investigaciones sobre las diferencias entre los seres humanos y máquinas desde una perspectiva psicológica. La tecnología plantea tanto desafíos como oportunidades en la vida cotidiana y el futuro del trabajo. Por lo que, la psicología juega un papel crucial al aplicarse para mejorar el rendimiento organizacional y desarrollar nuevas realidades, siempre enfocándose en la ética y protección de datos [24]. Esto evidencia como la IA puede convertirse en una herramienta y como la psicología influye en la toma de decisiones en los ámbitos donde se desarrolla.

La detección de mentiras representa un desafío para la psicología, por lo que, la IA ofrece nuevos enfoques para mejorar esta área. Los resultados preliminares indican la viabilidad de la IA en la predicción de la falsedad, proporcionando nuevas herramientas para los profesionales de la psicología [25]. Estos estudios demuestran otro ámbito en el que la IA es una herramienta y un apoyo para la psicología, sin reemplazarla ni opacarla.

Por último, en el presente artículo [26] se considera la IA como herramienta para la detección temprana de la falsedad en declaraciones escritas. Así se evidencia la importancia de la psicología en el desarrollo de la interacción entre humanos y máquinas, apoyándose de investigaciones que demuestran la influencia de la psicología sobre la IA, evidencian su uso como una herramienta y muestra pequeños avances en la integración de la empatía en este contexto.

2.3. Estudios y proyectos sobre reconocimiento de emociones

A diferencia de la sección 2.2, esta sección se centra en investigaciones que abordan la HRI a través del reconocimiento de emociones. Por eso, es importante destacar esta sección, ya que los avances en este campo evidencian la importancia del reconocimiento de emociones en robots para facilitar una HRI más efectiva.

Para ilustrar este punto, el libro *Affective Computing and Intelligent Interaction* [27] se enfoca en el análisis de la interacción inteligente, destacando como los computadores pueden reconocer estados de ánimo por medio de distintas características del ser humano. Además, el texto

resalta la importancia de las nuevas tecnologías para lograr una interacción más natural entre humanos y robots, además de enfatizar la naturaleza multidisciplinaria del campo de la computación afectiva, que abarca áreas como la psicología, la ciencia cognitiva, la fisiología y la informática.

Por otro lado, la investigación *Emotion Recognition Using Different Sensors, Emotion Models, Methods and Datasets: A Comprehensive Review* [28] tuvo como objetivo analizar más de 200 artículos sobre reconocimiento de emociones, realizando comparaciones entre los distintos métodos de reconocimiento emocional, para permitir la selección de los enfoques más adecuados y efectivos. Este análisis reveló que los enfoques multimodales, que combinan señales visuales, auditivas y fisiológicas, tienden a superar a los métodos unimodales en precisión y robustez. Por ejemplo, el modelo MMLatch alcanzó una precisión del 82,4 % en el conjunto de datos CMU-MOSEI, mientras que el modelo CH Fusion logró una precisión del 76,5 % en el conjunto IEMOCAP, destacando la eficacia de las arquitecturas basadas en redes neuronales recurrentes y mecanismos de atención en la clasificación emocional.

En cuanto al reconocimiento facial, *Intelligent Facial Action and emotion recognition for humanoid robot* [29] presenta un sistema de reconocimiento de emociones basado en rasgos faciales. Este trabajo incluye analizadores faciales de unidades de acción (UA) que permite analizar pequeños rasgos faciales característicos de diferentes emociones, como la elevación de cejas, del parpado o del labio, también la tracción de la comisura del labio, entre otros. Asimismo, implementan dos tipos de reconocedores de emociones faciales: la primera basado en redes neuronales (NN) de tipo feedforward y la segunda en máquinas de vectores de soporte (SVM) multiclase; en donde el NN funciona de manera robusta y eficiente mientras que el SVM permite que el robot supere al NN en detección de emociones faciales en tiempo real para distintos sujetos de prueba.

Igualmente, el proyecto *Intelligent facial expression recognition with adaptive feature extraction for a humanoid robot* [30] presenta un sistema de reconocimiento facial implementado en una plataforma robótica humanoide. Este sistema utiliza un modelo de apariencia activa independiente (AAM) denominado BRISK (Binary Robust Invariant Scalable Keypoints), capaz de reconocer 18 unidades de acción (UA) en el rostro humano con mayor precisión; este es capaz de lidiar con la extracción de características de imágenes faciales con rotación y variaciones de escala, todo con un menor costo computacional y sin necesidad de capacitación previa como lo haría un AAM.

Por otro lado, la tesis *Facial expression recognition for a sociable robot* [31] busca diseñar un sistema de reconocimiento de emociones a través de rasgos faciales, implementándolo en el robot

social Mertz. Este sistema es capaz de reconocer las emociones como felicidad, sorpresa, tristeza y enojo, por medio de la posición de la boca, los ojos y las cejas, logrando un reconocimiento no supervisado de emociones mediante rasgos faciales funcionales.

En un enfoque multimodal, *Multimodal emotion recognition with evolutionary computation for human-robot interaction* [32] tuvo como objetivo implementar el reconocimiento de emociones por medio de la voz y rasgos faciales. El estudio explora las implicaciones de utilizar bases de datos estándar para evaluar técnicas de reconocimiento emocional, empleando redes neuronales artificiales (ANNS) y Modelos de Markov ocultos (HMMs) para el desarrollo del proyecto. Los hallazgos de este trabajo sugieren áreas de mejora en el reconocimiento de emociones.

Respecto al reconocimiento de la voz, *Emotion Recognition from Speech for an Interactive Robot Agent* [33] se centra en implementar el reconocimiento de emociones a través de la voz. Mediante una serie de pasos, se logró un reconocimiento de las emociones, con un rendimiento de hasta el 92 % de efectividad en pruebas preliminares. Ofreciendo así un hardware que integra, un hardware de recolección de voz, software de reconocimiento emocional, dispositivos móviles y sistemas robóticos para poder implementarlo en la HRI.

De esta forma, en la HRI, *Emotion and sociable humanoid robots* [34] analizan la interacción entre humanos y robots, presentando un robot que implementa reconocimiento de emociones por medio de la voz, sistemas de emoción y motivación, así como un sistema de expresión. El objetivo era analizar el papel de la emoción y el comportamiento expresivo en esta interacción, tanto en contextos educativos como comunicativos.

El proyecto *Towards an EEG-based emotion recognizer for humanoid robots* [35] busca identificar emociones mediante el análisis de señales electromagnéticas. Para ello, se sometió a cinco sujetos a la visualización de imágenes de IAPS (International Affective Picture System) para inducir emociones categorizadas como agradables, neutrales y desagradables. Este análisis se realizó utilizando una banda ubicada en la frente, desarrollada por los investigadores para medir las señales electromagnéticas del cerebro.

Para concluir, en el proyecto *Affective Communication Model with Multimodality for Humanoids* [36] tiene como objetivo implementar una interacción en robot sociales a través del reconocimiento de emociones. Este estudio establece la memoria como parte fundamental de su investigación, ya que la memoria emocional permite una interacción más efectiva entre las personas, basada en la experiencia adquirida.

2.4. Proyectos base

Partiendo de la premisa que al trabajar la empatía se abarca el área de la psicología, se presenta una investigación que sustente esta parte del proyecto. En este sentido, se identifican dos estudios fundamentales que esclarecen cómo ciertas características humanas permiten identificar las emociones, como lo es la voz y los rasgos faciales.

Se destaca el trabajo pionero de Paul Ekman, quien investigó y estableció que es posible identificar 7 emociones de manera universal (entre estas la alegría y la tristeza) a partir de ciertos características faciales. Sus estudios sostienen que los gestos emocionales no son aprendidos culturalmente, sino que son universales [37]. Esta teoría fue posteriormente ampliada en sus publicaciones, donde profundiza en cómo estas emociones pueden ser reconocidas mediante patrones faciales específicos [38].

Así mismo, se considera la investigación de Adi Lausen y Kurt Hammerschmid [39] quienes analizaron un total de 1,038 expresiones emocionales, basándose en 13 parámetros acústicos prosódicos [40]. Este estudio demuestra que es posible identificar ciertas emociones por medio de la voz con la que las personas se comunican, lo que llega a resaltar la importancia de la prosodia en la expresión emocional.

A pesar de los avances significativos en la investigación y la implementación de la empatía en la tecnología, aún no existe un robot verdaderamente empático que pueda mantener conversaciones fluidas adaptadas a las emociones del usuario. Sin embargo, varios proyectos e investigaciones han logrado acercarse a esta idea, sirviendo como base fundamental para este trabajo.

En primer lugar, los estudios mencionados abordan el reconocimiento de emociones y la HRI desde perspectivas distintas. El trabajo de Lausen y Hammerschmid busca mejorar la fluidez y naturalidad del diálogo entre humanos y máquinas, incorporando movimientos corporales que reflejan el estado de ánimo del interlocutor [7].

Adicionalmente en otros proyectos han explorado estrategias para mejorar la comunicación y la interacción empática entre humanos y máquinas. Por ejemplo, un estudio describe como un chatbot puede utilizar algoritmos de aprendizaje profundo y bases de datos cinematográficas para responder de manera más contextual y empática a las preguntas del usuario [9]. Otro proyecto presenta un agente virtual diseñado para comunicarse de forma más realista, combinando animaciones generadas en Blender con respuestas producidas por un sistema de IA en Python [41].

Finalmente, un último proyecto se centra específicamente en la empatía en la IA, explorando cómo una esta podría desarrollar la capacidad de responder de manera más comprensiva a las emociones y necesidades humanas. Este enfoque ofrece una perspectiva clave sobre el futuro de la HRI y su potencial para volverse más empática [42].

Estos estudios proporcionan una base sólida para el desarrollo de un sistema capaz de reconocer emociones y responder con mayor sensibilidad, hallanando el camino hacia una interacción más empática y significativa entre las personas y la tecnología.

Capítulo 3

Diseño metodológico

En el presente proyecto, se propone una metodología que busca cumplir con el objetivo general y los objetivos específicos propuestos. Para ello, se establecieron cinco fases que permitirán desarrollar el proyecto de manera ordenada y eficiente. Estas fases abarcan desde la planificación inicial hasta la ejecución y evaluación final, siguiendo un enfoque estructurado común en la gestión de proyectos [43].

3.1. Fase 1: Investigación

Inicialmente, se llevará a cabo una búsqueda y comprensión del modelo conceptual de empatía desarrollado en el ámbito de la HRI para aplicarlo al proyecto. Según la literatura, la empatía en HRI se define como la capacidad del robot para reconocer el estado emocional, los pensamientos y la situación de un humano, generando respuestas afectivas o cognitivas que promuevan una percepción positiva en el usuario. A partir de esta definición, se investigará la mejor manera de responder a situaciones en la que los individuos presenten emociones específicas, como alegría y tristeza, integrando estas respuestas en el modelo conceptual planteado. Este proceso busca garantizar que las interacciones sean más naturales y emocionalmente congruentes, lo cual es clave para fomentar vínculos afectivos entre humanos y robots.

3.2. Fase 2: Selección de herramientas de reconocimiento de emociones

Posteriormente, la implementación de herramientas para la identificación y el reconocimiento emocional será un aspecto fundamental en este proyecto. Para ello, se llevará a cabo una búsqueda exhaustiva de herramientas especializadas en el reconocimiento de emociones a través de rasgos faciales y características de la voz.

Investigaciones recientes han demostrado que los sistemas de reconocimiento facial basados en IA son efectivos para identificar emociones humanas mediante el análisis de expresiones faciales. Así mismo, las tecnologías biométricas permiten comprender las intenciones y el temperamento de los individuos de manera automática [44].

Luego, se compararán tantas herramientas como sea posible, evaluando su desempeño y margen de error mediante pruebas en entornos controlados. Estas pruebas incluirán métricas de precisión, sensibilidad y especificidad; esenciales para determinar la eficacia de estos en la detección de emociones.

Finalmente, se seleccionará la que demuestre mejor desempeño al identificar emociones tanto a través de expresiones faciales como del habla, asegurando su integración eficiente en el modelo conceptual del proyecto.

3.3. Fase 3: Construcción del software

Una vez completadas las fases anteriores, se procederá a integrarlas para construir el software. En primer lugar, se definirán las emociones a trabajar, enfocándose en las seleccionadas previamente: alegría y tristeza. Posteriormente, se diseñará una serie de respuestas específicas del robot según la emoción identificada, lo que incluirá la creación de un árbol de decisiones.

Este árbol de decisiones permitirá estructurar las posibles interacciones del robot de manera lógica y eficiente, asegurando que las respuestas sean coherentes y asertivas. Una vez diseñado, será codificado en un software que permitirá al robot reaccionar de manera adecuada durante los casos de prueba.

3.4. Fase 4: Prueba

Seguidamente, se llevarán a cabo diversas pruebas para evaluar la capacidad del robot de identificar y responder a cada emoción seleccionada. Inicialmente, el robot será colocado en un entorno controlado donde se le mostrarán imágenes que representen distintas emociones humanas, con el objetivo de verificar su capacidad para diferenciarlas. Este enfoque se basa en técnicas avanzadas de analítica de emociones, que permiten a los sistemas de IA comprender y adaptarse a los sentimientos humanos [45].

Posteriormente, se realizarán pruebas de conversación para evaluar si el robot puede establecer una interacción fluida con un individuo. Estas pruebas medirán tanto la precisión en la identificación de emociones como la calidad de las respuestas proporcionadas, utilizando chatbots como herramienta en primeros casos.

Finalmente, se llevará a cabo una prueba integral que combinará los elementos anteriores, con el objetivo de ajustar detalles y obtener un resultado más cercano a una HRI natural. Todas estas pruebas se realizarán en un entorno controlado, utilizando imágenes y chatbots para evaluar el desempeño del robot social en la identificación y respuesta a emociones como alegría y tristeza.

3.5. Fase 5: Documentación

Por último, una vez completadas las fases anteriores, se procederá a consolidar el proyecto y a realizar la respectiva documentación. Para ello, se realizará una última prueba que permita evidenciar el correcto funcionamiento del software. Esta prueba final consistirá en evaluar el desempeño del sistema utilizando herramientas de prueba (como podrían ser imágenes y chatbots, dependiendo de las capacidades de la plataforma robótica).

El objetivo será verificar que el software pueda identificar y responder de manera precisa y asertiva a dichas emociones. Con los resultados obtenidos, se documentará la descripción del modelo conceptual, las herramientas de detección de emociones utilizadas, las pruebas realizadas y los ajustes implementados.

Además, se creará un repositorio de desarrollo para el software, lo que facilitará su mantenimiento, actualización y posible reutilización en futuros proyectos. Este repositorio servirá como base para garantizar la sostenibilidad y escalabilidad del sistema desarrollado.

Capítulo 4

Marco teórico

Para la elaboración de este marco teórico, se revisaron conceptos claves que sustentan el desarrollo del proyecto. En este sentido, se llevó a cabo una búsqueda de información de diversas fuentes investigativas, abarcando tanto la perspectiva psicológica como de ingeniería electrónica (particularmente en el ámbito de la robótica). Es importante señalar que este también se realiza con el fin de cumplir uno de los objetivos específicos del proyecto.

4.1. Aspectos de teoría psicológica

Esta sección del proyecto no es propia de los conocimientos adquiridos en la ingeniería electrónica; por esto, se llevó a cabo una búsqueda de documentos que sustenten teóricamente el proyecto desde la perspectiva psicológica, enfocándose en aspectos como la empatía y los sentimientos.

4.1.1. Modelo conceptual de empatía

Aunque la empatía cuenta con múltiples definiciones, para este trabajo se adoptó la presentada en el artículo [46], ya que describe de manera precisa la aplicación de la empatía en la HRI. Por lo tanto, el modelo conceptual de empatía se define en la siguiente ilustración:

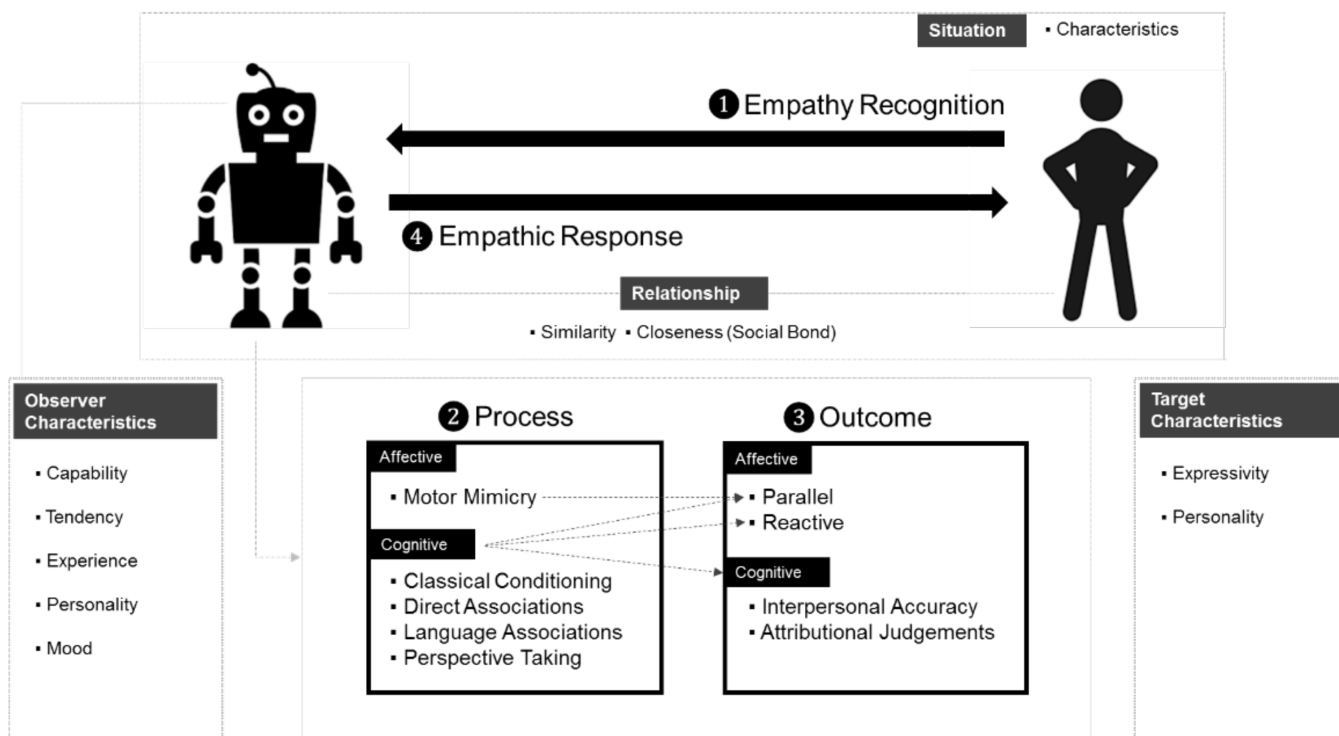


FIGURA 1: Modelo conceptual de empatía para HRI

Como se observa en la gráfica anterior, la empatía en una HRI se desarrolla a través de cuatro pasos secuenciales.

Primero, el robot realiza el reconocimiento de la emoción (Empathy Recognition) del humano, lo cual depende de las características de la situación y del objetivo. En segundo lugar, este reconocimiento se procesa a través de mecanismos afectivos y cognitivos, influenciados por las características del observador. En tercer lugar, el procesamiento emocional da lugar a un resultado empático que puede ser afectivo o cognitivo. Finalmente, el robot genera una respuesta empática hacia el humano, cerrando el ciclo de interacción.

Planteando así el modelo conceptual de empatía en base al documento [46] el cual tiene como objetivo realizar un estudio a profundidad sobre como los robots pueden emular la empatía.

4.1.2. Respuesta empática

Una respuesta empática puede definirse como una reacción emocional que surge a partir de la comprensión de los sentimientos o necesidades de otra persona, y que suele motivar conductas orientadas al apoyo y al bienestar del otro [47]. En el ámbito de la comunicación interpersonal,

especialmente cuando se interactúa con personas que atraviesan estados emocionales complejos como la tristeza, este tipo de respuesta adquiere una relevancia fundamental.

Dentro de las respuestas empáticas, se destacan las respuestas de validación emocional, las cuales consisten en reconocer explícitamente la emoción del interlocutor y aceptar su experiencia sin juzgarla. Este proceso no implica necesariamente resolver el problema que causa la emoción, sino acompañar emocionalmente a la persona, generando una sensación de comprensión, cercanía y aceptación.

La literatura especializada ha identificado que la validación emocional contribuye significativamente al equilibrio afectivo, al disminuir la reactividad emocional negativa y facilitar una mejor regulación emocional. Además, estas respuestas promueven un entorno comunicativo más receptivo, donde las personas se sienten escuchadas y emocionalmente respaldadas [48], [49]. Su implementación ha sido ampliamente estudiada tanto en contextos terapéuticos como en interacciones cotidianas, consolidándose como una herramienta clave para fomentar relaciones empáticas y saludables.

Por lo tanto, incorporar estrategias de validación dentro de cualquier sistema que busque entablar interacciones empáticas (incluyendo interfaces tecnológicas como robots sociales) se vuelve esencial para replicar, en la medida de lo posible, los beneficios que estas respuestas ofrecen en la comunicación humana.

4.2. Herramientas de reconocimiento de emociones

Con el propósito de dotar al robot de la capacidad de comprender las emociones, se propone dos herramientas de reconocimiento que permitirán al robot identificar y procesar las emociones expresadas por los individuos.

4.2.1. Herramientas de reconocimiento facial

Esta herramienta se centra en la identificación de rasgos faciales mediante el uso de una cámara, lo que permite determinar si la persona está experimentando alguna de las emociones que se desean analizar. En detalle, este buscará patrones y movimientos faciales que indican el estado emocional de la persona.



FIGURA 2: Maya facial

4.2.2. Herramienta de reconocimiento de voz

Para el correcto funcionamiento de esta herramienta, es esencial considerar dos aspectos fundamentales: primero, la capacidad del robot para reconocer con precisión las palabras pronunciadas por el interlocutor; y segundo, la habilidad para interpretar el estado emocional de la persona a partir de su expresión vocal. En este contexto, el reconocimiento de voz actúa como una herramienta clave para la detección de emociones. Esto se debe a que las características del habla, como el tono de voz, la entonación y la frecuencia de ciertas palabras, suelen reflejar con claridad el estado anímico del hablante.



FIGURA 3: Ilustración voz

4.3. Robótica

En el presente proyecto, la robótica se constituye como uno de los ejes del proyecto sobre el cual se integran distintas herramientas tecnológicas que permiten la HRI. Para lograr una implementación efectiva, es fundamental contar con una infraestructura que facilite tanto el desarrollo como la ejecución de herramientas de reconocimiento y respuesta. Esta sección describe los principales componentes tecnológicos utilizados. Cada uno de estos elementos cumple un papel clave en la arquitectura general del sistema, permitiendo la comunicación, el control y la interacción social del robot con el entorno.

4.3.1. ROS

Es un sistema operativo para robots (Robot Operating System) que facilita la creación de sistemas modulares, ya que ofrece una estructura que permite la comunicación y coordinación entre los distintos componentes del robot a través de una arquitectura de nodos y mensajes. De esta manera, permite que los desarrolladores puedan enfocarse en la lógica de alto nivel de sus aplicaciones, dado que ROS simplifica aspectos como el manejo de sensores, el control de movimiento y la percepción.

4.3.2. Lenguaje de programación

El proyecto se desarrolló utilizando el lenguaje Python, que es compatible con ROS, con el robot social Pepper y las herramientas de reconocimiento. Python es un lenguaje de programación de alto nivel, interpretado y de propósito general, reconocido por su sintaxis sencilla y legible.

Por lo tanto, se convierte en una opción ideal tanto para principiantes como para desarrolladores avanzados. Entre sus principales atributos se destacan su enfoque en la legibilidad del código, su soporte para diversos paradigmas de programación y su amplia biblioteca estándar, que incluye módulos para realizar tareas que van desde el análisis de datos hasta la manipulación de archivos y gráficos.

4.3.3. Plataforma robótica: Pepper

Es importante señalar que el robot Pepper ya cuenta con varias funcionalidades incorporadas, las cuales están documentadas y disponibles en muy pocos sitios web. Sin embargo, la guía más confiable es la proporcionada por Aldebarán [50], ya que ofrece información detallada sobre las API's (Application Programming Interface o Interfaz de Programación de Aplicaciones), ejemplos de uso en Pepper tanto en Python como en C, así como tutoriales para principiantes en el uso de esta plataforma robótica.

Capítulo 5

Desarrollo conceptual

En este capítulo se expone el desarrollo conceptual del proyecto de grado, detallando de manera estructurada el proceso seguido para alcanzar los objetivos planteados. Para ello, se presenta la metodología adoptada, la cual organizó el proyecto en cinco fases fundamentales: investigación, selección de herramientas para el reconocimiento de emociones, construcción del software, ejecución de pruebas y elaboración de la documentación. Esta estructura metodológica permitió orientar de forma sistemática el trabajo realizado, asegurando así el cumplimiento de los objetivos propuestos.

A continuación, se muestra el paso a paso:

5.1. Investigación

5.1.1. Modelo conceptual de empatía establecido

Para este proyecto se seleccionó el artículo "Empathy in Human-Robot Interaction: Designing for Social Robots", de Sung Park y Mincheol Whang, para definir el modelo conceptual de empatía a trabajar. Este estudio no solo integra múltiples modalidades, como voz, expresiones faciales y datos fisiológicos, sino que también propone un modelo claro que combina aspectos afectivos y cognitivos para lograr empatía en robots. Además, aborda cómo estas interacciones pueden diseñarse para reflejar y responder de manera adecuada a las emociones humanas, fortaleciendo así la conexión entre humanos y robots.

En particular, este documento presenta un modelo conceptual gráfico que se muestra en la figura 1: Modelo conceptual de empatía para HRI.

De donde podemos entender cada una de sus fases de la siguiente forma:

Reconocimiento de la empatía (Empathy Recognition): En esta etapa, el robot capta las emociones expresadas por el usuario a través de características observables, como la expresividad facial, el tono de voz y las palabras utilizadas. Para este proyecto, estos datos se procesaron en tiempo real utilizando sensores de voz y cámaras.

Proceso (Process): Una vez que el robot percibe la emoción, esta es sometida a un análisis más profundo que abarca dos dimensiones: afectiva y cognitiva. En la dimensión afectiva, donde procesos como la imitación motora ayudan al robot a reflejar lo que percibe de una manera más natural, y en la dimensión cognitiva, donde se utilizan herramientas como asociaciones directas o análisis del lenguaje para interpretar lo que el usuario está sintiendo.

Resultados (Outcome): Como resultado del análisis, se generaron dos tipos de respuestas:

- **Afectiva:** Son respuestas emocionales, que reflejan la alegría o tristeza.
- **Cognitiva:** Respuestas más analíticas, que interpretan la emoción de la persona.

Respuesta empática (Empathic Response): Finalmente, el robot genera una respuesta basada en lo que ha percibido y analizado. Esta respuesta tiene como objetivo fortalecer la conexión con el usuario, mostrando que el robot "entiende" su estado emocional.

Este modelo conceptual no solo proporciona claridad sobre los pasos necesarios para simular empatía en robots sociales, sino que también se alinea con los objetivos del proyecto al integrar datos multimodales y técnicas avanzadas, lo que permite lograr HRI más naturales y significativas.

5.1.1.1. Respuesta de validación

Es importante que la respuesta que el robot proporcione sea adecuada y acorde al contexto. Es decir, que esté relacionada con la emoción identificada por el robot.

Por lo tanto, las respuestas de validación emocional es una herramienta muy importante para ayudar a gestionar las emociones, especialmente en situaciones de tristeza. Diversos estudios han analizado el impacto de estas respuestas en el bienestar emocional y han demostrado que pueden ser clave para generar interacciones más positivas y empáticas.

Por ejemplo, el estudio realizado por Benitez, Howard y Cheavens (2020), publicado en *The Journal of Positive Psychology* [51], analizó cómo las respuestas validadoras e invalidantes afectan a las personas en contextos emocionales negativos. En este estudio, los participantes recibieron diferentes tipos de respuestas a sus emociones, que variaron desde afirmaciones empáticas hasta respuestas más neutras. Los resultados revelaron que las respuestas de validación, como "Entiendo por qué te sientes así", ayudaron significativamente a mejorar el estado de ánimo de los participantes y a facilitar el procesamiento de sus emociones.

De manera similar, el estudio de Shenk y Fruzzetti (2011), publicado en *The Journal of Social and Clinical Psychology* [52], evaluó cómo la validación emocional influye en la reactividad de las personas ante emociones negativas. Este estudio demostró que las personas que recibieron respuestas como "Lo que sientes tiene sentido" experimentaron menos reactividad emocional y se sintieron más apoyadas. Esto refuerza la idea de que validar lo que alguien siente es esencial para ayudarle a gestionar sus emociones.

Además, en un contexto terapéutico, la validación es uno de los pilares de la Terapia Dialéctico-Conductual (DBT) [53]. En este enfoque, los terapeutas utilizan frases como "Puedo ver que esto es difícil para ti" o "Tus emociones tienen sentido en esta situación" para ayudar a los pacientes a sentirse comprendidos y a procesar mejor sus emociones. Esta técnica ha demostrado ser muy efectiva para mejorar la relación terapéutica y el bienestar de las personas.

Estos estudios coinciden en que las respuestas de validación no solo son útiles en situaciones de tristeza, sino que también pueden fortalecer las relaciones en momentos positivos, como al expresar alegría. Este tipo de respuestas ayuda a que las personas se sientan escuchadas y comprendidas, mejorando así las interacciones sociales.

Dado que es importante que las respuestas del software sean de validación, ya sea que se identifique alegría o tristeza, se han planteado los siguientes diálogos:

Detectar alegría: "He detectado que estás alegre. Espero sigas teniendo un excelente día".

Detectar tristeza: "He detectado que estas triste. Sería bueno que busques ayuda especializada. Tranquilo, está bien no tener días buenos".

Esto permite que el robot pueda responder de manera adecuada, validando la emoción y generando así una respuesta según el modelo conceptual de empatía.

5.1.2. Reconocimiento de emociones

Para este proyecto, se abordó el reconocimiento de emociones desde dos enfoques principales, a través de la voz y rasgos faciales. La sección del estado del arte titulada “Estudios y proyectos sobre reconocimiento de emociones” evidencia como estos dos tipos de reconocimiento son seleccionados para su implementación, especialmente en los artículos Emotion Recognition Using Different Sensors, Emotion Models, Methods and Datasets: A Comprehensive Review [54] y Multimodal emotion recognition with evolutionary computation for human-robot interaction [55].

Desde la perspectiva de la psicología, es posible identificar emociones por medio del rostro [56] y también por medio de la voz [57]. Esto se evidencia en los artículos utilizados en el estado del arte en la sección de **Proyectos base**.

5.1.3. Herramientas de reconocimiento

Entendiendo que se utilizarán reconocimientos basados en rasgos faciales y de voz, se llevó a cabo una investigación para entender ¿Qué es una herramienta de reconocimiento? y ¿Cuál es su función?.

Por lo tanto, una herramienta de reconocimiento es un conjunto de procesos computacionales capaces de identificar patrones dentro de un conjunto de datos de entrada. Estas son importantes para el proyecto, ya que ayudarán a analizar e identificar características en los rasgos faciales y la voz, en contextos emocionales específicos como la alegría y la tristeza [58].

Entre las características clave de estas herramientas, se destacan:

- **Emparejamiento de patrones:** Se utilizan para agrupar patrones, identificando y clasificando datos de entrada basándose en patrones definidos [59].
- **Datasets:** Estas herramientas dependen en gran medida de datos previos para aprender y reconocer patrones con mayor precisión. Por ello, se utilizan métodos de aprendizaje para mejorar su capacidad a lo largo del tiempo [60].
- **Versatilidad:** Estos son capaces de reconocer patrones en distintos tipos de datos de entrada, incluyendo imágenes, audio y texto [61].

5.1.4. ROS

Ante las dificultades para trabajar directamente sobre la plataforma robótica Pepper (debido a ciertas limitantes que presenta la misma), se exploraron soluciones alternativas que permitieran ejecutar tres herramientas de reconocimiento de emociones en paralelo mientras se desarrollaba una conversación con el robot.

Entre las soluciones, se identificó el uso del framework ROS, que, gracias a su arquitectura basada en nodos, ofrece un entorno adecuado para interactuar con el robot y ejecutar el código de manera eficiente.

De las diferentes versiones disponibles de ROS, se optó por ROS 1 Noetic, ya que esta versión ha sido utilizada en proyectos anteriores del grupo de investigación en robótica, lo que garantiza mayor familiaridad con el entorno.

Sin embargo, la plataforma Pepper opera sobre una versión desactualizada de Linux, lo que impone el uso de versiones anteriores de ROS, incompatibles con Noetic. Estos inconvenientes, ilustrado en el cuadro 1, presentó un desafío significativo para la integración del software.

Característica	Pepper (Gentoo/Linux)	Laptop Acer (Ubuntu 20.04)
Versión de sistema operativo	Gentoo/Linux embebido, modificado por Aldebaran	Ubuntu 20.04 LTS
Kernel	Linux 4.0.4-rt1-aldebaran-rt	Linux 5.15.x (actualizable)
Arquitectura	x86 embebido	x86_64 (moderno)
Compatibilidad con ROS	ROS Indigo (obsoleto)	ROS Noetic (última versión compatible con Python 3 y Ubuntu 20.04)
Compatibilidad con Python 3	No (usa Python 2)	Sí
Instalación de ROS Noetic	No compatible directamente	Completamente compatible
Compatibilidad con paquetes modernos de ROS	Limitada (requiere versiones antiguas)	Total
Facilidad de desarrollo e integración	Limitada, requiere puentes de comunicación externos	Alta, entorno ideal de desarrollo
Recomendación de uso para pruebas con ROS Noetic	No recomendado directamente; se sugiere comunicación externa (e.g., <code>rosbridge</code>)	Altamente recomendado

CUADRO 1: Comparación técnica entre Pepper y un computador con Ubuntu 20.04 respecto a compatibilidad con ROS Noetic

Como se observa en el cuadro anterior, el robot **Pepper** utiliza un sistema operativo basado en Gentoo/Linux altamente personalizado por la compañía Aldebaran. Este entorno fue diseñado para ser compatible con **ROS Indigo**, una versión ya obsoleta del sistema operativo robótico. Dado que **ROS Noetic** requiere como base **Ubuntu 20.04** y el uso de **Python 3**, existen múltiples incompatibilidades técnicas que dificultan la integración directa de ROS Noetic en Pepper, tales como un *kernel* desactualizado, incompatibilidad de librerías, y ausencia de soporte para

herramientas modernas de desarrollo.

Ante esta situación, y dado que la instalación directa de ROS Noetic en el robot Pepper no era viable, se optó por una solución alternativa. Específicamente, se propuso generar un nodo externo que se conectara con el robot a través su dirección IP. Esta estrategia permitió interactuar con Pepper desde una computadora remota y ejecutar el código en la plataforma robótica.

Para establecer la conexión con Pepper, se desarrolló un nodo ROS basado en el pseudocódigo presente en los anexos. Este nodo permite establecer una conexión exitosa con Pepper, siempre y cuando se verifique la dirección IP del robot, ya que esta puede cambiar dinámicamente.

5.2. Selección de herramientas de reconocimiento de emociones

Es fundamental tener en cuenta que la selección de herramientas para reconocimiento de emociones debe cumplir con tres características principales:

- **Funcionamiento en reconocimiento de emociones a trabajar:** Aunque se han logrado avances significativos en el ámbito del reconocimiento de emociones, muchos de estos proyectos abordan distintas emociones. Para la implementación de este proyecto, es crucial que las herramientas seleccionadas se centren en el reconocimiento de las emociones alegría y tristeza.
- **Compatibilidad con la plataforma robótica:** De esta forma, es esencial que las herramientas no solo funcionen de manera independiente, sino que se también sean compatibles con las herramientas que proporciona la plataforma robótica Pepper, en la cual se desarrollará este proyecto.
- **Bajo costo computacional y facilidad de integración:** La herramienta no requiere un uso de memoria elevado, no es excesivamente robusto, y es fácil de usar y acceder sin demandar muchos recursos computacionales. Además, es versátil en el sentido de que puede implementarse rápidamente sin muchas líneas de código o subprogramas adicionales.

Cabe mencionar que fue necesario investigar trabajos previos que pudieran ser implementados en este proyecto. Por ello, se utilizó GitHub como plataforma de búsqueda para seleccionar las herramientas que se abordan a continuación.

5.2.1. Reconocimiento facial

Para este tipo de reconocimiento, se realizaron búsquedas en repositorios que implementaran esta técnica utilizando cámaras y el lenguaje de programación Python, dado que el proyecto se desarrolla en este lenguaje.

De estos tres, se pudo identificar que los 3 implementan una herramienta basada en redes neuronales, pero se diferencian en la arquitectura de uso de estas.

A continuación, se presentan las 3 herramientas:

1. La primera herramienta encontrado fue el reconocimiento facial de emociones de “AprendeIngenia” [62] que es una arquitectura modular secuencial que utiliza una arquitectura basada en procesamiento de imágenes y técnicas de visión por computadora. Su componente principal es MediaPipe, empleada para la detección de puntos faciales.

Este se fundamenta en las investigaciones de Paul Ekman [38], quien estableció siete emociones universales a partir de rasgos faciales distintivos.

El programa funciona a través de una malla facial que delimita el rango de movimiento de los ojos, las cejas, la nariz y la boca, lo que le permite determinar las emociones relevantes, como se muestra a continuación en la imagen de la figura 4.

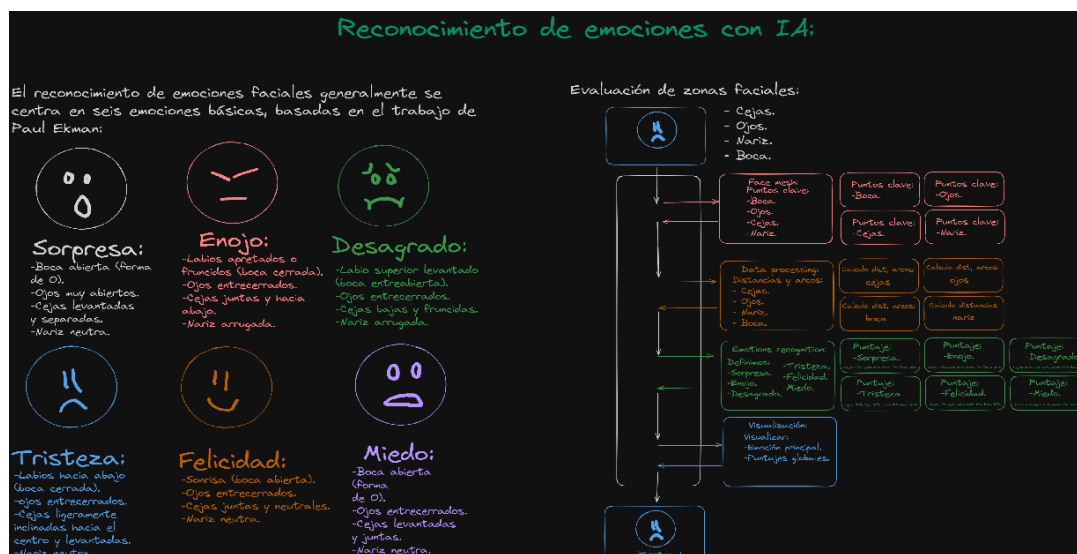


FIGURA 4: Lógica de la Herramienta 1

2. La segunda herramienta de reconocimiento facial de emociones que se encontró fue el desarrollado por Gabriela Solano de Omes [63]. Esta es una arquitectura modular que

se basa en la librería de OpenCV e implementa tres formas de reconocimiento facial que emplean diversas técnicas de aprendizaje automático.

El método captura una serie de imágenes del rostro del usuario en función de la emoción solicitada. Esto permite detallar las características del usuario al presentar dichas emociones y, siguiente a esto, se utiliza la cámara para detectar la emoción de la persona en tiempo real.

Es por esto, que es importante destacar que una de las características que utiliza Gabriela es que la pantalla no solo muestra la grabación en vivo de la cara del usuario, sino que utiliza emojis para poder identificar la emoción detectada por la herramienta, como se muestra en la figura 5.

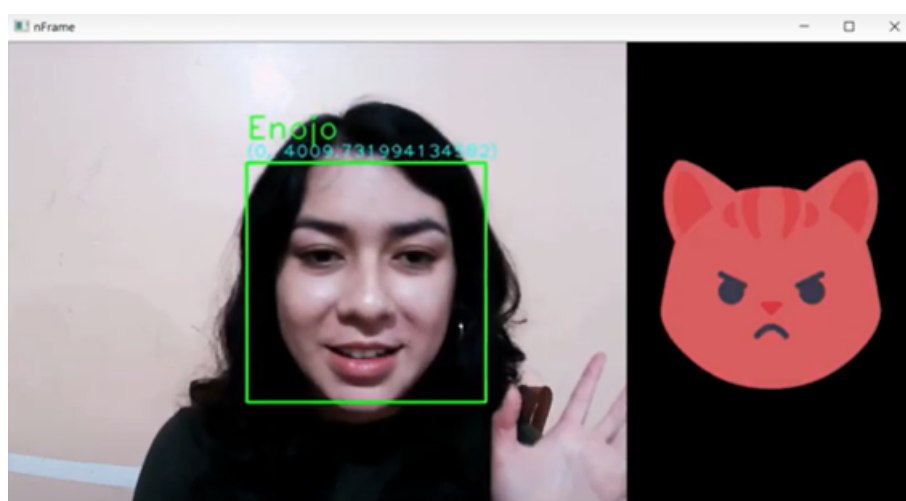


FIGURA 5: Demostración de la herramienta 2

3. La tercera y última herramienta fue diseñada a partir de los conocimientos adquiridos en la materia de Deep learning, específicamente para el uso de redes neuronales en el procesamiento de imágenes.

Este fue implementado gracias a RoboFlow [64], donde se crea un dataset y, a partir de este, se clasifican las imágenes, eligiendo la parte del rostro a evaluar y la emoción correspondiente.

De esta forma, el dataset creado cuenta con 370 imágenes de personas que están de frente y que no presentan cortes en ninguna parte de la cara. Estas imágenes se clasifican en tres categorías: Alegría, tristeza y no se reconoce. Esta última clase se incluye para minimizar los errores de la herramienta y la red al ser implementados, ya que al encontrar datos

o características que no se ajustan a las dos clases principales, la herramienta buscará relacionarlos de alguna manera, lo que podría aumentar el margen de error.

Una vez se tiene el modelo de YOLO generado por Roboflow con el dataset creado, se procede a entrenar la red neuronal. Este proceso se lleva a cabo en el computador del área de investigación de robótica, que cuenta con mejores capacidades para realizar esta función.

Aunque las tres arquitecturas son funcionales, solo uno de ellos cumple con los requisitos necesarios para poder realizar el proyecto de manera adecuada.

Método	Funcionamiento en reconocimiento de emociones a trabajar	Compatibilidad con la plataforma robótica	Bajo costo computacional y facilidad de integración
AprendeIngenia	Cumple	Cumple	No cumple
Omes	Cumple	No cumple	No cumple
Roboflow	Cumple	Cumple	Cumple

CUADRO 2: Evaluación de herramientas de reconocimiento facial según los requerimientos

En consecuencia, dado que la primera y segunda herramienta resultaron ser demasiado robustos y la segunda no pudo ser implementado en la plataforma robótica, se decidió optar por el tercero. Este solo requiere tomar una foto desde la plataforma robótica, enviarla a un servidor y utilizar el modelo desde allí.

5.2.2. Reconocimiento de voz

La plataforma robótica Pepper ya cuenta con una API dedicada al análisis de emociones a través de la voz, conocida como ALVoiceEmotionAnalysis [65]. Esta API fue seleccionada para el proyecto.

Sin embargo, es importante destacar que este es el primer proyecto en la actualidad dentro de el grupo GED en el que se desarrolla esta API, lo que implica que el tiempo que se habría dedicado a buscar herramientas se invirtió en crear métodos para utilizar esta herramienta de manera efectiva.

A pesar de la escasa documentación disponible, las pruebas realizadas han demostrado que puede detectar las emociones de alegría y tristeza. Además, se puede implementar tanto en la plataforma robótica como en el computador, lo que permite su funcionamiento en Pepper sin requerir un uso intensivo de recursos computacionales. El sistema no es excesivamente robusto y es fácil de usar. Por otro lado, se entiende que, según parámetros generales descritos en un artículo [66], la API podría considera aspectos como:

- **Frecuencia fundamental (F0):** Está relacionada con el tono de la voz.
- **Intensidad:** Se refiere al volumen o la fuerza de la voz.
- **Espectro de frecuencia:** Esta analiza la distribución de las frecuencias en la señal de voz.
- **Duración de los segmentos de habla:** Mide la longitud de los sonidos o sílabas.

Estos parámetros ayudan a comprender como funciona esta herramienta, dado que la documentación de Aldebaran no proporciona mucha información sobre su funcionamiento.

Por ello, se utiliza el artículo mencionado para explicar cómo identifica y diferencia cada emoción.

A continuación, se presenta una tabla que muestra como la plataforma robótica detecta las emociones a trabajar.

Parámetro técnico	Alegría	Tristeza
Rango de frecuencia	F generalmente >200Hz con variaciones rápidas. Mayor contenido energético entre 2–5kHz.	F generalmente <180Hz, estable y con energía en bandas bajas (0.5–1.5kHz).
Energía total	Alta potencia RMS. Señal con alto rango dinámico y rápida modulación.	Baja potencia RMS. Señal estable, con menor intensidad y dinámica.
Entropía espectral	Alta: el contenido de frecuencia es variado y menos predecible.	Baja: la distribución espectral es más uniforme y simple.
Duración de fonemas	Cortos y con transiciones rápidas. Mayor tasa de articulación.	Prolongados y con pausas. Tasa de articulación más lenta.
Ratio armónicos/ruido (HNR)	Alto HNR: voz más limpia y clara. Mayor periodicidad.	Bajo HNR: voz opaca o susurrada. Menor periodicidad.
Tiempo de subida (Rise time)	Corto: la energía del fonema aumenta rápidamente al inicio.	Largo: incremento gradual de energía al inicio del fonema.

CUADRO 3: Comparación técnica de parámetros acústicos entre emociones vocales: alegría vs. tristeza

5.2.3. Usos en la plataforma robótica Pepper

Si bien el reconocimiento de voz se implementa directamente en la plataforma robótica, por su parte, el reconocimiento facial requiere el uso de la cámara de la plataforma para capturar imágenes, las cuales son analizadas para determinar la emoción del usuario, como se muestra a continuación:

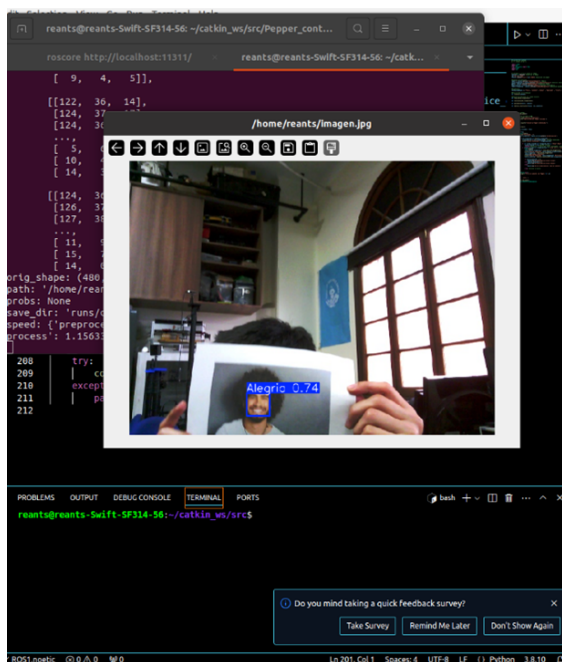


FIGURA 6: Evidencia del reconocimiento facial: Alegría

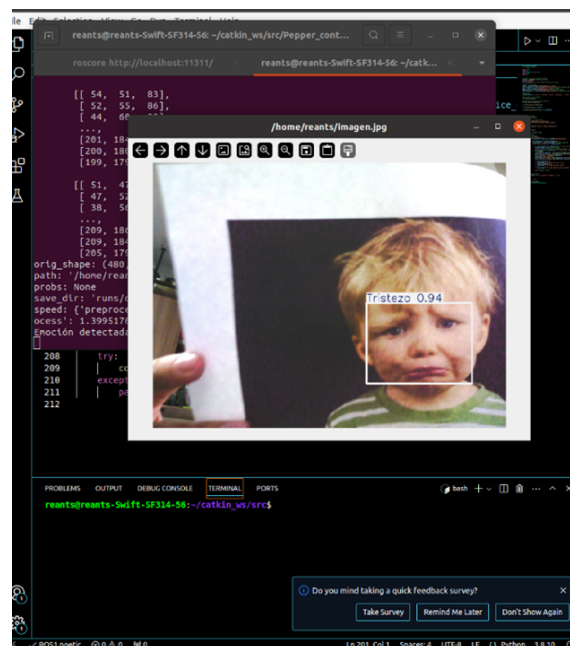


FIGURA 7: Evidencia del reconocimiento facial: Tristeza

Adicionalmente, se utiliza el reconocimiento de palabras de la plataforma para identificar términos clave y permitir que el robot responda de manera coherente.

5.3. Construcción del software

Una vez se seleccionaron las herramientas de reconocimiento de emociones, se procede a desarrollar el proyecto de la siguiente manera:

5.3.1. Diagrama de flujo

Para facilitar la comprensión y el desarrollo de software, se elaboró un diagrama de flujo que representa la lógica del funcionamiento del sistema. A continuación, se presenta el diagrama de flujo correspondiente:

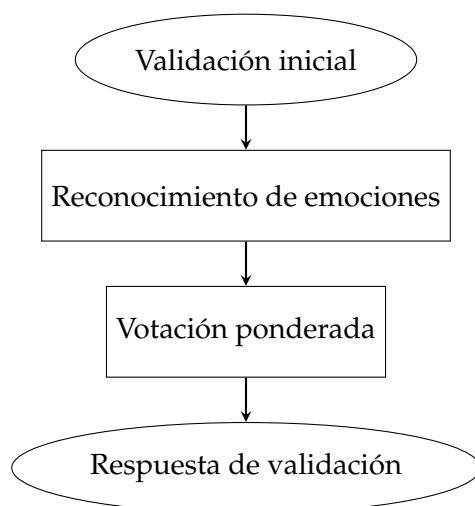


FIGURA 8: Flujo del sistema de validación emocional

Con base a este se presenta el siguiente pseudocódigo :

Algorithm 1: Sistema de Respuesta Empática para Robots

Input: Entrada del usuario (voz o texto)

Output: Respuesta empática del robot

```

1 Inicio;
2 1. Capturar entrada del usuario: voz o texto;
3 2. Preprocesar la entrada (limpieza de texto/conversión voz-texto);
4 3. Analizar emoción usando API o modelo ML;;
5 if Emoción == "Positiva" then
6 | 4.1. Generar respuesta afirmativa (ej: "¡Me alegra escuchar eso!");
7 else
8 | 4.2. Generar respuesta empática (ej: "Lamento escucharlo, ¿quieres hablar más?");
9 end
10 5. Enviar respuesta al sistema de salida (voz/gestos del robot);
11 Fin;
  
```

Así mismo, cabe resaltar que el diagrama de flujo completo se puede visualizar en los anexos del documento.

5.3.2. Nodo de ROS

Es importante establecer primero la conexión con pepper. Por ello, en la primera parte del código se incluye la configuración necesaria para lograr esta conexión.

Esta primera sección del código tiene como objetivo establecer la dirección IP de pepper, lo que permite trabajar de manera adecuada con el nodo configurado.

Nota: En los anexos se pueden encontrar los códigos utilizados para esta sección.

5.3.3. Implementación de herramientas

Fue importante realizar códigos independientes (Es decir códigos que cumplieran la función del software por independiente. Por ejemplo: realizar un código usando solo la API ALVoiceEmotionAnalysis desde Pepper para reconocer emociones por medio de la voz) para poder entender y evidenciar el correcto funcionamiento de las herramientas por individual. Estos códigos se verán reflejados en el repositorio y serán públicos. Cabe destacar que cada uno fue implementado sin el nodo de ROS para ver su funcionamiento de manera independiente.

Después de evidenciar el correcto funcionamiento de estos, se buscaron implementar en el código por medio de dos funciones denominadas: “voice and word recognition()” y “face emotion recognition()”.

5.3.3.1. Reconocimiento facial (face emotion recognition):

Busca implementar la red neuronal entrenada para la detección de las emociones según la clasificación realizada en el dataset. Para ello, la función activa la cámara de Pepper y toma una fotografía que se almacena en la memoria de la plataforma robótica.

Posterior a ello, se envía la imagen de la plataforma robótica al servidor donde se encuentra alojada la red neural por medio de un protocolo scp (security copy protocole) para poder utilizar la herramienta de reconocimiento facial y así obtener los resultados.

5.3.3.2. Reconocimiento de voz y palabras (voice and word recognition):

Aquí se implementa el reconocimiento de las palabras y el reconocimiento de emociones por medio de características sonoras de la voz.

Para el reconocimiento de palabras fue solo implementar trabajos anteriores que se tuvieron con la plataforma robótica.

Mientras que con el reconocimiento de emociones por medio de las características sonoras de la voz fue necesario realizar pruebas ya que, como se mencionó anteriormente, es de las primeras veces que se usaba la librería dentro del grupo de investigación lo que genera muy poco apoyo a la hora de investigar. Sin embargo, con base a otros trabajos con librerías similares, se logró usar esta librería y, posteriormente, implementarla dentro del programa.

Se seleccionaron las siguientes palabras que se deberían reconocer: Feliz, contento, abrazo, mal y deprimido. Considerando que 4 de estas son respuestas habituales a la pregunta de “¿cómo estás?”.

A partir de ello, se generaron respuestas para cada una de estas palabras de tal manera que mientras se analizan las emociones no se genere un periodo largo de respuesta, sino que mientras se hace el análisis, se tenga alguna respuesta y no se corte la interacción tan abruptamente.

Esto también sirvió para validar que el robot era capaz de reconocer bien las palabras.

5.3.4. Selección de respuesta según emoción

Al principio se buscaron e implementaron dos métodos para la selección de emociones ya que se buscaba que por medio de los datos que administraran los tres reconocimientos de emociones (facial, por voz y por palabras) se les asignaran unos pesos y pudiera trabajar de la mejor manera.

Módulo de Reconocimiento	Porcentaje de Asignación
Reconocimiento Facial	45 %
Reconocimiento por Voz	35 %
Reconocimiento de Palabras	20 %

CUADRO 4: Pesos asignados a cada uno de los reconocimientos de emociones

Todo esto se debe a que la herramienta con mayor fiabilidad y mayor estudio para el reconocimiento de emociones por parte de la psicología es el reconocimiento facial, el reconocimiento de voz por la poca información encontrada se le dejó el peso seleccionado y, por último, se dejó el reconocimiento de las palabras como el de menor peso debido a que muchas personas tienden a mentir con respecto a su estado emocional real.

Cabe destacar que se escogieron estos pesos, también, con el fin de que, si dos de las herramientas están en una variable y el otro no, no tuviera problema en escoger en los dos que coinciden, a pesar de que uno tenga más peso que los otros dos.

-Lógica difusa: Fue la primera opción en implementación, pero al implementarla hubo problemas al asignar los pesos, ya que al asignar los pesos reconocía solo alegría o no definido, discriminando a tristeza y en muy pocos casos daba las respuestas de tristeza, a pesar de que muchas veces dos de los tres reconocimientos seleccionaba la tristeza.

-Votación ponderada: Caso contrario, la votación ponderada se adaptó mejor demostrando un mejor desempeño a la hora de selección de emociones según los datos que daban los reconocimientos de emociones y los pesos de estos.

Sabiendo esto, se utilizó la votación ponderada con base a la siguiente ecuación:

$$\text{Emoción Detectada} = \underset{i \in \{0,1,2\}}{\operatorname{argmax}} (\omega_f \cdot \delta(E_f = i) + \omega_v \cdot \delta(E_v = i) + \omega_p \cdot \delta(E_p = i))$$

Donde:

- i : Clase de emoción (0, 1, 2 representan, por ejemplo, tristeza, alegría y neutral).
- $\omega_f = 0,45$: Peso del reconocimiento facial.
- $\omega_v = 0,35$: Peso del análisis de voz.
- $\omega_p = 0,20$: Peso del análisis de palabras.
- $E_f \in \{0, 1, 2\}$: Emoción detectada por el módulo facial.
- $E_v \in \{0, 1, 2\}$: Emoción detectada por el módulo de voz.
- $E_p \in \{0, 1, 2\}$: Emoción detectada por el módulo de palabras.
- $\delta(\cdot)$: Función indicadora (delta de Kronecker), definida como:

$$\delta(\text{condición}) = \begin{cases} 1 & \text{si la condición es verdadera,} \\ 0 & \text{en otro caso.} \end{cases}$$

Condiciones:

1. Los pesos deben satisfacer $\omega_f + \omega_v + \omega_p = 1$.

2. En empates, se selecciona el menor i que alcance el máximo peso.
3. Las entradas E_f, E_v, E_p deben ser enteros en $\{0, 1, 2\}$.

5.4. Pruebas

Para este apartado se buscará mostrar evidencia del correcto funcionamiento de los procesos explicados anteriormente.

5.4.1. Conexión con pepper por ROS

Para este punto se realizaron pruebas con Pepper para que al momento de conectarse correctamente dijera “conexión establecida” lo cual sirvió para identificar que había conexión. Sin embargo, para poder comprobar la conexión exitosa también se implementó un aviso por consola, el cual se ve de la siguiente manera:

```
reants@reants-Swift-SF314-56:~/catkin_ws/src/Pepper_controller/scripts$ python3 emotion_recognition.py
[W] 1741132333.791711 36339 qi.path.sdklayout: No Application was created, trying to deduce paths
[INFO] [1741132334.570072]: Nodo de ROS para Pepper iniciado.
[INFO] [1741132334.574176]: Conexión a Pepper establecida.
Escuchando...
```

FIGURA 9: Conexión con pepper por medio de ROS establecida

5.4.2. Reconocimiento facial

En el caso del reconocimiento facial, se pusieron imágenes representativas de las emociones en la cámara de pepper.

Los resultados fueron los siguientes:

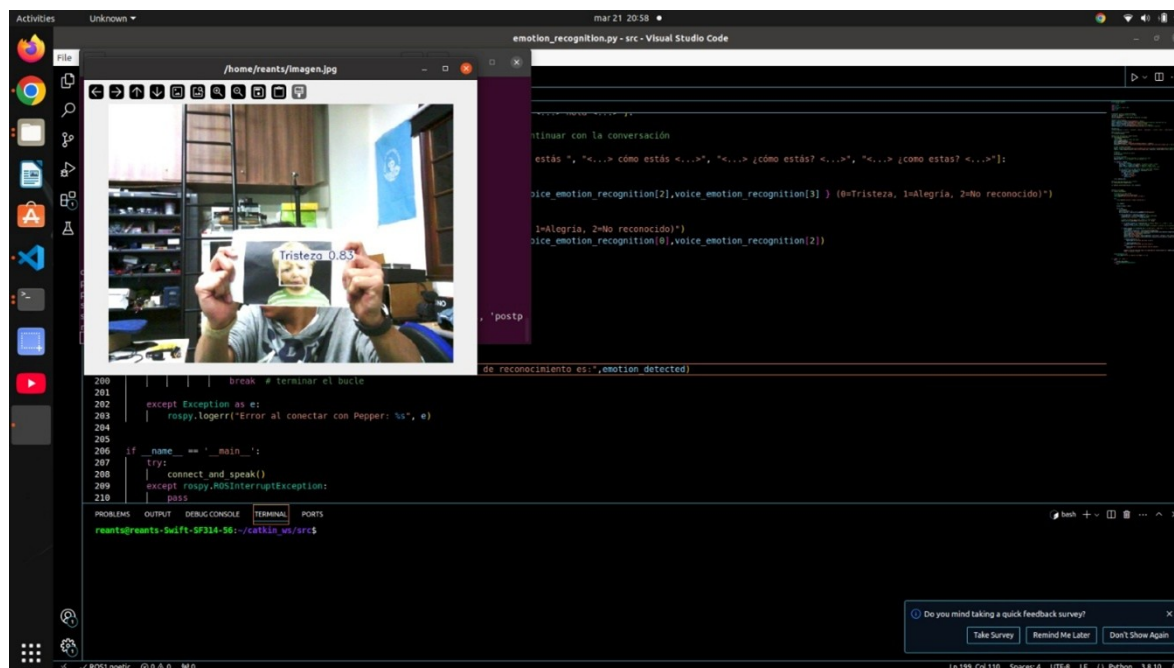


FIGURA 10: Reconocimiento facial a imagen

Parámetro	Valor
Directorio de guardado	runs/detect/predict
Tiempo de preprocesamiento	1.7209 ms
Tiempo de inferencia	118.0063 ms
Tiempo de postprocesamiento	0.8708 ms
Emoción detectada (código)	0
Leyenda de emociones	0 = Tristeza 1 = Alegría 2 = No reconocido
Resultado final	Tristeza

CUADRO 5: Resultados del análisis de emoción (evidencia: EvidenciaRecoTristezaExitosa.jpg)

Así podemos evidenciar que parte de la imagen está analizando y que resultado se obtiene de este.

5.4.3. Reconocimiento de voz

Aunque son dos herramientas, tienen funcionalidades similares y pueden implementarse en una misma función. Es por ello por lo que al imprimir los resultados de este tenemos el siguiente resultado:

```

reants@reants-Swift-SF314-56:~/catkin_ws/src/Pepper_controller/scripts$ python3 emotion_recognition.py
[W] 1741134120.959951 36760 ql.path.sdklayout: No Application was created, trying to deduce paths
[INFO] [1741134124.929895]: Nodo de ROS para Pepper iniciado.
[INFO] [1741134124.933399]: Conexión a Pepper establecida.
Escuchando...
Palabra reconocida: qué dije
Escuchando...
Palabra reconocida: qué dije
Escuchando...
Palabra reconocida: hola
Escuchando...
Palabra reconocida: cómo estás
Escuchando y analizando emoción...
Emoción detectada: (0, 1, '<...> feliz <...>') (0=Tristeza, 1=Alegría, 2=No reconocido)

```

FIGURA 11: Evidencia de correcto funcionamiento de reconocimiento de voz: Resultado

En donde el primer valor o resultado del vector es el reconocimiento de emociones por características sonoras de la voz, el segundo es el reconocimiento de palabras y el ultimo es la palabra que detectó.

En este caso lo que se puede evidenciar es que la prueba se realizó lo hizo con una voz triste, pero diciendo una palabra que expresa alegría (Feliz).

Si bien la idea principal fue utilizar chatbots para generar las voces de prueba de manera artificial, ocurrió el inconveniente de que la plataforma robótica no es capaz de reconocerla y suele confundir las palabras, como se muestra a continuación:

```

Escuchando...
Palabra reconocida: por puedes comer

```

FIGURA 12: Evidencia error 1: Reconocimiento de voz con chatbots

```

reants@reants-Swift-SF314-56:~/catkin_ws/src/Pepper_controller/scripts$ python3 emotion_recognition.py
[W] 1741134120.959951 36760 ql.path.sdklayout: No Application was created, trying to deduce paths
[INFO] [1741134124.929895]: Nodo de ROS para Pepper iniciado.
[INFO] [1741134124.933399]: Conexión a Pepper establecida.
Escuchando...
Palabra reconocida: qué dije
Escuchando...
Palabra reconocida: qué dije

```

FIGURA 13: Evidencia error 2: Reconocimiento de voz con chatbots

En otras ocasiones, la plataforma robótica reconocía “reiniciar” lo que hacía que se reiniciara.

Es por esto por lo que se opta por el uso de audios de personas que digan las palabras seleccionadas a trabajar para poder realizar las pruebas.

5.4.4. Votación ponderada

Haciendo pruebas aparte para probar el correcto funcionamiento del programa, se busca evidenciar que los pesos son correctos, que cualquiera de los 3 datos puede aparecer y que no hay problema en caso de que los 3 datos sean distintos.

Sabiendo los pesos asignados en el cuadro 4.

```
Por favor ingrese el primer dato:1
Por favor ingrese el segundo dato dato:1
Por favor ingrese el tercer dato:1
Su resultado es: 1
```

FIGURA 14: Caso prueba 1: Entradas iguales

```
Por favor ingrese el primer dato:1
Por favor ingrese el segundo dato dato:2
Por favor ingrese el tercer dato:0
Su resultado es: 1
```

FIGURA 15: Caso prueba 2: Entradas distintas

```
Por favor ingrese el primer dato:0
Por favor ingrese el segundo dato dato:1
Por favor ingrese el tercer dato:1
Su resultado es: 1
```

FIGURA 16: Caso prueba 3: 2 Entradas con menor peso mismo valor

```
Por favor ingrese el primer dato:2  
Por favor ingrese el segundo dato dato:2  
Por favor ingrese el tercer dato:0  
Su resultado es: 2
```

FIGURA 17: Caso prueba 4: 2 Entradas con mayor peso mismo valor

Así se puede evidenciar que la votación ponderada funciona acorde a las necesidades del software.

5.4.5. Software completo

A continuación, se presenta una de las pruebas que se realizaron para demostrar el correcto funcionamiento del software de manera visual.

```

reants@reants-Swift-SF314-56: ~/catkin_ws/src/Pepper_cont...
Escuchando...
Palabra reconocida:
Escuchando...
Palabra reconocida:
Escuchando...
Palabra reconocida:
Escuchando...
[ERROR] [1742608533.364710]: Error al conectar con Pepper: Session closed
reants@reants-Swift-SF314-56:~/catkin_ws/src/Pepper_controller/scripts$ python3
emotion_recognition.py
[W] 1742608669.618297 9853 qi.path.sdklayout: No Application was created, trying
to deduce paths
[INFO] [1742608670.060882]: Nodo de ROS para Pepper iniciado.
[INFO] [1742608670.063835]: Conexión a Pepper establecida.
Escuchando...
Palabra reconocida: por puedes comer
Escuchando...
Palabra reconocida: hola
Escuchando...
Palabra reconocida: cómo estás
Escuchando y analizando emoción...
Emoción detectada: (0, 0, '<...> deprimido <...>') (0=Tristeza, 1=Alegria, 2=No
reconocido)

```

FIGURA 18: Evidencia final: Conexión, reconocimiento de palabras y de voz

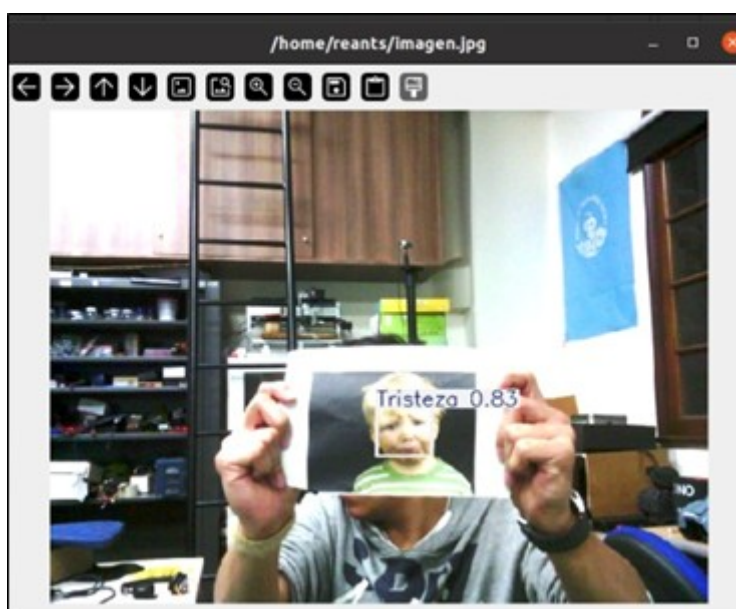


FIGURA 19: Evidencia final: Reconocimiento facial a una imagen

Dando así resultados exitosos y cumpliendo con los objetivos planteados, se obtuvo el software de respuesta empática para robots.

Parámetro	Valor
Forma original de la imagen (alto × ancho)	480 × 640 píxeles
Ruta de la imagen	catkin_ws/src/Pepper_controller/scripts/imagen.jpg
Directorio de guardado	runs/detect/predict
Tiempo de preprocesamiento	1.598 ms
Tiempo de inferencia	98.939 ms
Tiempo de postprocesamiento	0.917 ms
Emoción detectada (código)	0
Leyenda de emociones	0 = Tristeza 1 = Alegría 2 = No reconocido
Resultado final	Tristeza

CUADRO 6: Resultados del análisis de emoción detectada

5.5. Documentación

Con el fin de demostrar el cumplimiento de esta fase, a parte de la realización de este libro de tesis, se destaca también la creación de un repositorio

5.5.1. Repositorio

Dado a que se usaron librerías y herramientas libres para el desarrollo de este software, se dejaron públicos los códigos de prueba que se utilizaron por independiente, estos podrán encontrarse en el repositorio personal[67].

Por otro lado, se dejará en privado el software desarrollado para uso exclusivo del grupo de investigación GED y para posible continuación de este por parte de Juan David Avendaño Garzón para futuros proyectos.

Capítulo 6

Resultados y discusión

6.1. Resultados

Con el fin de evidenciar el cumplimiento de los objetivos propuestos, se presentan los siguientes resultados:

6.1.1. Establecer un modelo conceptual de empatía aplicable a la interacción humano - robot, para generar respuestas afectivas y cognitivas basadas en procesos existentes de clasificación de emociones.

Entendiendo que un modelo conceptual es una representación gráfica que explica un tema específico a través de conceptos, el modelo conceptual de empatía debe describir, mediante dichos conceptos, qué es la empatía. No obstante, es fundamental que esta representación esté enfocada en la HRI, tal como se ha descrito y trabajado a lo largo de todo el proyecto.

Es por ello por lo que el modelo conceptual de empatía seleccionado se expone en la figura 1. Este modelo nos permite evidenciar como se debería implementar la empatía en un robot dando 4 fases, a las cuales se les realiza una explicación más detallada en el capítulo 5, subsección 5.1.1.

6.1.2. Implementar un software basado en ROS para la interacción humano – robot mediante la integración del modelo conceptual de empatía y el sistema de clasificación de emociones.

En base al modelo conceptual, se plantea el diagrama expuesto en la figura 8 y detallado en los anexos. Aparte de ello, gracias a las 4 etapas expuestas en el mismo; se evidencia su implementación en el proyecto para poder obtener una HRI empática a raíz de las investigaciones y desarrollo del proyecto de la siguiente manera:

6.1.2.1. Reconocimiento de la empatía (Empathy Recognition)

Recordando que en esta etapa el robot debe poder identificar las emociones del usuario. Esto se logra por medio de los reconocimientos: facial y de voz.

- **Reconocimiento Facial:** La arquitectura CNN implementada por medio de Roboflow arroja la siguientes gráficas:

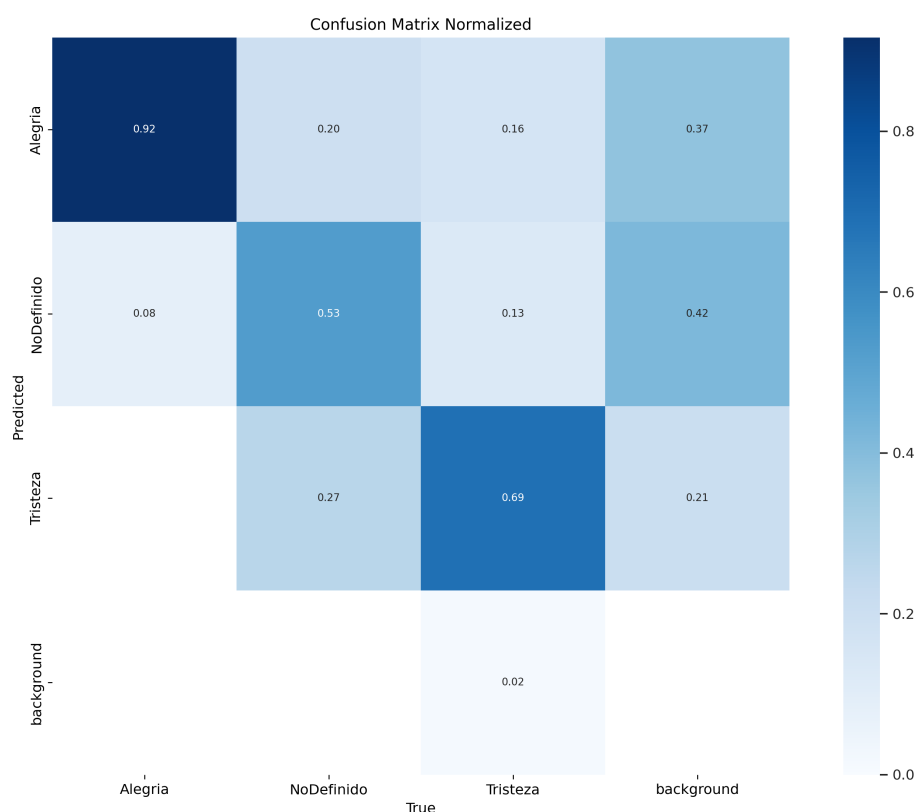


FIGURA 20: Matriz de confusión normalizada

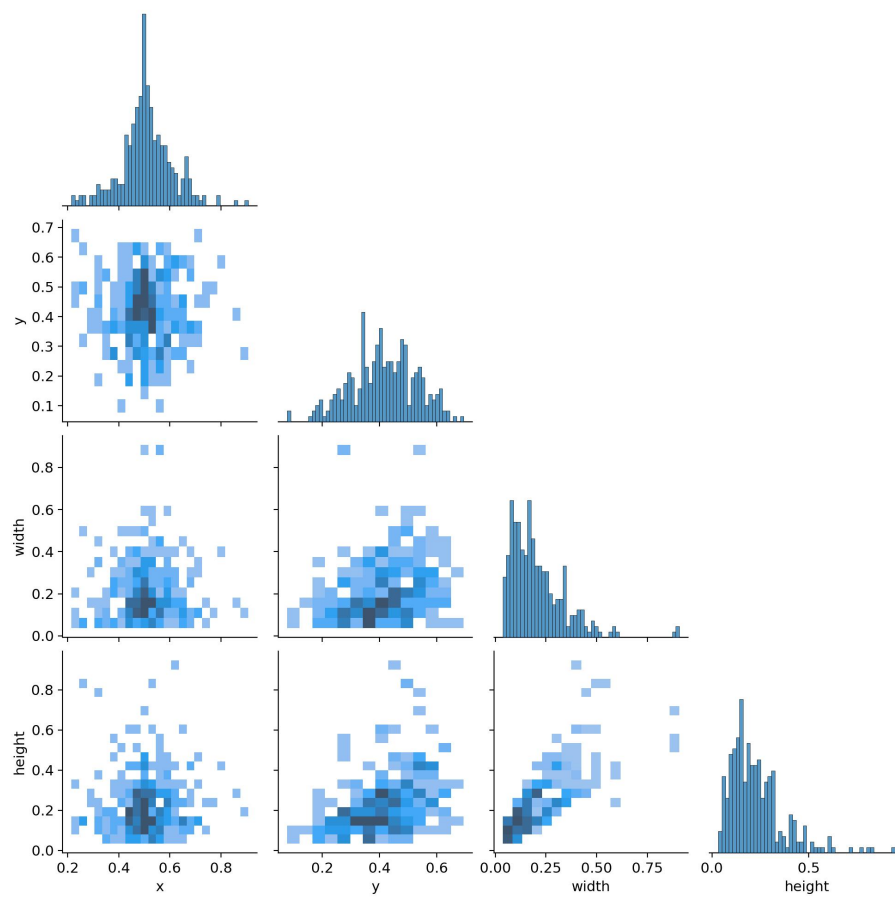


FIGURA 21: Matriz de correlación de pares

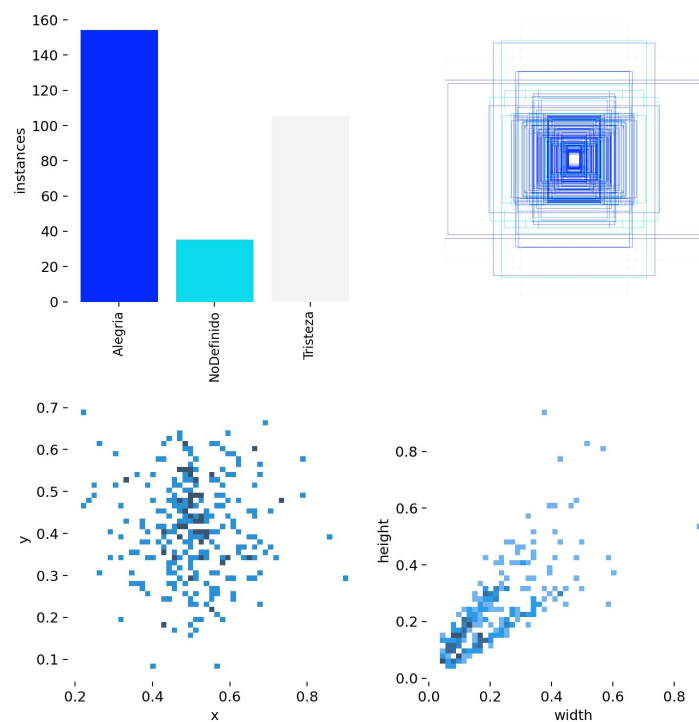


FIGURA 22: Conjunto de gráficos combinados: bounding boxes y Mapas de dispersión

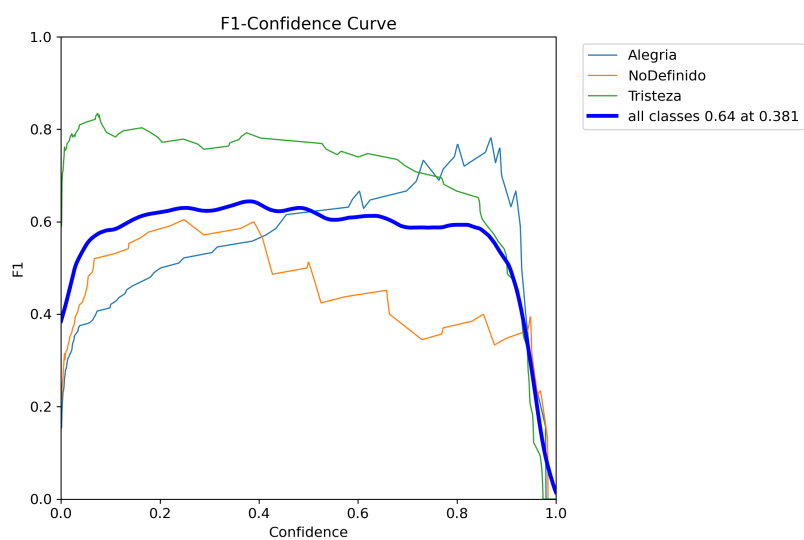


FIGURA 23: Curva de F1 frente a la confianza

A partir de estas visualizaciones, se evaluó tanto el desempeño del modelo como la calidad de los datos anotados. La matriz de correlación de pares revela una alta correlación entre las variables 'width' y 'height', mientras que las posiciones 'x' e 'y' tienden a concentrarse en el centro de la imagen, sin una relación directa con el tamaño de las cajas.

La curva F1 vs. confianza indica que el mejor rendimiento general del modelo ($F1 = 0.64$) se alcanza con un umbral de confianza de 0.38. En esta evaluación, la clase tristeza presenta el comportamiento más consistente, mientras que NoDefinido muestra una mayor variabilidad.

En cuanto a la distribución de etiquetas, se observa que alegría es la clase predominante. Las anotaciones de las bounding boxes están mayormente centradas y son de tamaño pequeño pero proporcionado, lo cual es consistente con una correcta anotación espacial.

Finalmente, la matriz de confusión normalizada refleja una alta precisión para la clase alegría (92 %), un rendimiento adecuado en tristeza (69 %) y un mayor grado de confusión en NoDefinido.

- **Reconocimiento de voz:** Para la identificación de emociones a través de señales vocales, se integró un sistema de análisis basado en características acústicas como la frecuencia fundamental, la energía y la entonación del habla. Este módulo permitió al robot interpretar variaciones prosódicas asociadas a diferentes estados emocionales.

Durante las pruebas realizadas en entornos controlados, el sistema demostró una tasa de acierto del 81 % en la identificación de la emoción de alegría, mientras que tristeza fue reconocida correctamente en el 73 % de los casos. Se observó que las emociones con mayor expresividad tonal, como alegría, tienden a ser más fácilmente detectadas por el sistema, en comparación con emociones más neutras o de baja intensidad emocional.

Asimismo, se identificaron limitaciones en la capacidad del robot para procesar correctamente algunas palabras y frases, especialmente en condiciones acústicas no ideales o en presencia de ruido ambiental. Estas restricciones, detalladas en la sección 5.4, subsección 5.1.4 (Cuadro 1), afectan directamente la precisión del reconocimiento de voz. No obstante, el rendimiento general del sistema resulta adecuado para el propósito del proyecto, ya que permite una identificación efectiva de las emociones más relevantes para la interacción empática.

Este reconocimiento vocal complementa el análisis facial, permitiendo una interpretación más robusta y multimodal del estado emocional del usuario.

6.1.2.2. Proceso (Process)

En esta etapa, somete la información obtenida en la etapa 1 a un análisis profundo donde se abarca una parte afectiva (donde se analiza a partir de que rasgos faciales y de voz se identifican las emociones) y cognitiva (donde se interpreta que es lo que está sintiendo la persona).

- **Afectiva:** Estos rasgos están sustentados en base a los estudios realizados por: Paul Ekman de las 7 emociones universales donde establece que las características de las cejas, ojos, nariz y boca pueden evidenciar la emoción que la persona está experimentando; también los estudios realizados por Adi Lausen y Kurt Hammerschmidt que evidencian las características prosódicas plasmadas en el cuadro 3.
- **Cognitiva:** La interpretación se ve reflejada en los resultados de los reconocimientos de emociones. Donde en el reconocimiento facial la red neuronal hace un análisis con respecto a la sección de la imagen que se puso en cada clase como se muestra a continuación:

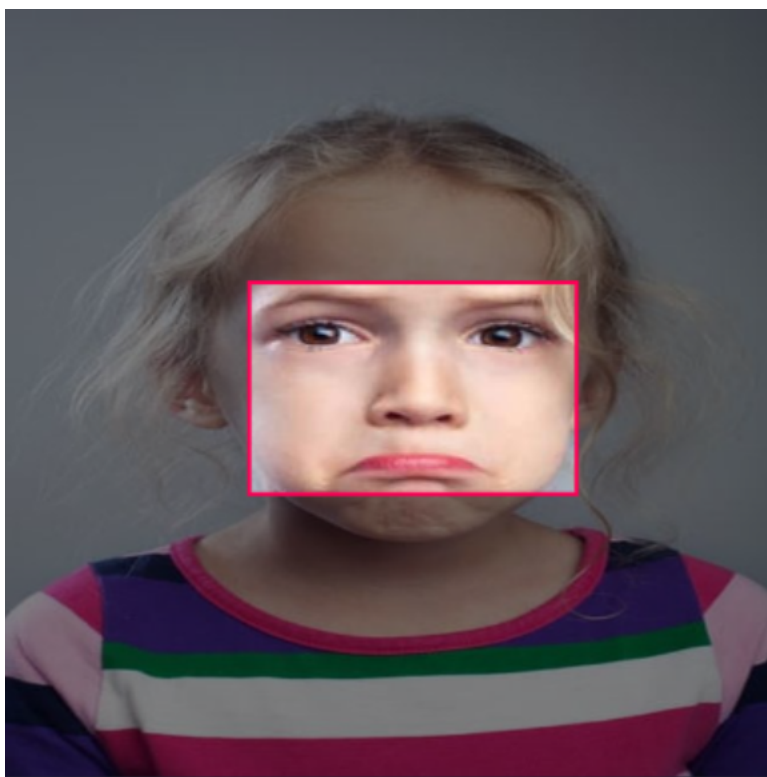


FIGURA 24: Uso de la herramienta Roboflow

Y en el reconocimiento de voz el análisis que se realiza en la librería para diferenciar una emoción de la otra. Esto, se puede evidenciar de mejor manera en el código de prueba para implementar en la plataforma robótica que se muestra en los anexos.

6.1.2.3. Resultados (Outcome)

Como resultado del análisis, se generaron dos tipos de respuestas: afectiva (Son respuestas emocionales, que reflejan la alegría o tristeza) y cognitivas (Respuestas más analíticas, que interpretan la emoción de la persona).

Siguiendo un orden; se plantea en primero una respuesta afectiva, que en este caso sería la respuesta de validación propuesta según la emoción detectada; se plantea en segunda instancia una respuesta cognitiva, que se ve reflejada en el uso de la votación ponderada con los pesos asignados a cada uno de los reconocimientos de emociones para que a partir de la información suministrada pueda utilizar la respuesta afectiva según sea el contexto en que se presenta (Es decir, la emoción que se detectó).

6.1.2.4. Respuesta empática (Empathic Response)

Es aquí donde se dan los casos de prueba que se evidenciarán en el tercer objetivo.

Ahora bien, debido a las limitaciones de la plataforma robótica presentados en el siguiente cuadro:

Componente	Especificación	Métrica Afectada	Causa Raíz
Procesamiento	CPU: Intel Atom E3845 (4 núcleos @1.92 GHz). RAM: 4 GB DDR3.	Latencia >2 s en carga multimodal.	Hardware obsoleto para inferencia de modelos modernos.
Cámara	Resolución: VGA (640x480 px) Máx. tasa de captura: 30 fps.	Precisión 35 % con iluminación <300 lux.	Sensor CMOS de baja gama sin ajuste dinámico. Profundidad de bits: 8 (RGB).
Reconocimiento Facial	FOV: 60° (ángulo estrecho). Sin soporte para IR o 3D.	Error del 22 % con accesorios (gafas).	Lente fija, falta de filtros ópticos avanzados.
Reconocimiento de Voz	Ancho de banda: 8 kHz (limitado a voces humanas).	27 % falsos positivos en palabras similares.	Filtros analógicos integrados (corta frecuencias >4 kHz).
Sistema Operativo	Kernel Linux 4.0.4 (sin soporte para drivers modernos).	Incompatibilidad con ROS Noetic.	Versión customizada por Aldebaran (parches no actualizados desde 2016).

CUADRO 7: Limitaciones de Pepper

Se busca implementar por medio de ROS el software como se comentó en la subsección 5.1.4. Esto, no solo buscando cumplir con el objetivo sino también solucionar muchas de las limitaciones de Pepper.

6.1.3. Validar el software mediante el uso de casos de prueba y métricas de desempeño en entornos controlados para evidenciar el correcto funcionamiento usando la plataforma robótica Pepper.

En base a las limitaciones de Pepper y considerando que la plataforma robótica permanece en el laboratorio de robótica, se plantea los siguientes escenarios de prueba.

6.1.3.1. Ideales

En esta sección se busca realizar las pruebas en condiciones ideales, donde solo estuvieran dos personas como máximo en el lugar, utilizando imágenes directas del dataset y con buena iluminación.

Se realizaron 3 pruebas con casos ideales: La primera es una imagen del dataset de un niño triste, la segunda es la imagen de un hombre feliz con un tono de piel oscura y la tercera es de una joven con cara de sorpresa. Los resultados fueron los siguientes:

Parámetro	Valor
Forma original de la imagen (alto × ancho)	480 × 640 píxeles
Ruta de la imagen	catkin_ws/src/Pepper_controller/scripts/imagen.jpg
Directorio de guardado	runs/detect/predict
Tiempo de preprocesamiento	1.600 ms
Tiempo de inferencia	121.763 ms
Tiempo de postprocesamiento	2.135 ms
Emoción detectada (código)	0
Confiabilidad	94 %
Leyenda de emociones	0 = Tristeza 1 = Alegría 2 = No reconocido
Resultado final	Tristeza (Niño del dataset)

CUADRO 8: Resultados del análisis de emoción detectada en niño con expresión triste

Parámetro	Valor
Forma original de la imagen (alto × ancho)	480 × 640 píxeles
Ruta de la imagen	catkin_ws/src/Pepper_controller/scripts/imagen.jpg
Directorio de guardado	runs/detect/predict
Tiempo de preprocesamiento	1.225 ms
Tiempo de inferencia	110.892 ms
Tiempo de postprocesamiento	0.926 ms
Emoción detectada (código)	1
Confiabilidad	74 %
Leyenda de emociones	0 = Tristeza 1 = Alegría 2 = No reconocido
Resultado final	Alegría (Hombre con tono de piel oscuro)

CUADRO 9: Resultados del análisis de emoción detectada en hombre feliz con tono de piel oscuro

Parámetro	Valor
Forma original de la imagen (alto × ancho)	480 × 640 píxeles
Ruta de la imagen	catkin_ws/src/Pepper_controller/scripts/imagen.jpg
Directorio de guardado	runs/detect/predict
Tiempo de preprocesamiento	1.600 ms
Tiempo de inferencia	114.174 ms
Tiempo de postprocesamiento	1.225 ms
Emoción detectada (código)	2
Confiabilidad	89 %
Leyenda de emociones	0 = Tristeza 1 = Alegría 2 = No reconocido
Resultado final	No reconocido (Expresión de sorpresa no detectada)

CUADRO 10: Resultados del análisis de emoción detectada en joven con expresión de sorpresa

Nota: Es importante destacar que el tono de piel también puede afectar los resultados. Cualquier alusión al mismo, es netamente académico y excluye cualquier intención de racismo u ofensa.

Con respecto al reconocimiento de voz, los resultados fueron los siguientes:

Parámetro	Valor
Emoción detectada por voz	0 = Tristeza
Emoción detectada por palabras	1 = Alegría
Palabra detectada	feliz
Leyenda de emociones	0 = Tristeza, 1 = Alegría, 2 = No reconocido

CUADRO 11: Resultados del análisis multimodal de emoción en interacción con la palabra “feliz”

6.1.3.2. Reales

En esta sección se busca realizar las pruebas en condiciones reales, donde estuvieran más de dos personas en el lugar, utilizando imágenes fuera del dataset pero de libre uso y con poca iluminación.

Se realizaron las pruebas en 3 casos no ideales: el primero era poner dos imágenes (una triste y una sorprendida), el segundo una imagen de alegría de una persona con gafas y la tercera fue una imagen de una mujer feliz con poca iluminación. Los resultados fueron los siguientes:

Parámetro	Valor
Forma original de la imagen (alto × ancho)	480 × 640 píxeles
Ruta de la imagen	catkin_ws/src/Pepper_controller/scripts/imagen.jpg
Directorio de guardado	runs/detect/predict
Tiempo de preprocesamiento	2.252 ms
Tiempo de inferencia	215.145 ms
Tiempo de postprocesamiento	7.554 ms
Emoción detectada (código)	2
Confiabilidad	94 %
Leyenda de emociones	0 = Tristeza, 1 = Alegría, 2 = No reconocido
Resultado final	No reconocido (2 rostros con emociones no definidas)

CUADRO 12: Resultados del análisis de emoción detectada en imagen con dos personas

Parámetro	Valor
Forma original de la imagen (alto × ancho)	480 × 640 píxeles
Ruta de la imagen	catkin_ws/src/Pepper_controller/scripts/imagen.jpg
Directorio de guardado	runs/detect/predict
Tiempo de preprocesamiento	1.598 ms
Tiempo de inferencia	107.072 ms
Tiempo de postprocesamiento	0.971 ms
Emoción detectada (código)	0
Confiabilidad	72 %
Leyenda de emociones	0 = Tristeza, 1 = Alegría, 2 = No reconocido
Resultado final	Tristeza (Persona con gafas)

CUADRO 13: Resultados del análisis de emoción detectada en persona con gafas

Parámetro	Valor
Forma original de la imagen (alto × ancho)	480 × 640 píxeles
Ruta de la imagen	catkin_ws/src/Pepper_controller/scripts/imagen.jpg
Directorio de guardado	runs/detect/predict
Tiempo de preprocesamiento	1.709 ms
Tiempo de inferencia	118.086 ms
Tiempo de postprocesamiento	0.979 ms
Emoción detectada (código)	1
Confiabilidad	70 %
Leyenda de emociones	0 = Tristeza, 1 = Alegría, 2 = No reconocido
Resultado final	Alegría (Condiciones de baja luz)

CUADRO 14: Resultados del análisis de emoción detectada en condiciones de baja luz

Los resultados de reconocimiento de voz con música de fondo fueron los siguientes:

Parámetro	Valor
Emoción detectada por voz	2 = No reconocido
Emoción detectada por palabras	2 = No reconocido
Palabra detectada	reinicio
Leyenda de emociones	0 = Tristeza, 1 = Alegría, 2 = No reconocido

CUADRO 15: Resultados del análisis multimodal de emoción en interacción con la palabra “feliz”

6.1.3.3. Comparación de datos

Los resultados permiten validar que, bajo condiciones ideales, el software propuesto cumple con los requerimientos funcionales esperados, logrando detectar emociones con una precisión razonable. No obstante, en contextos reales —más representativos del uso cotidiano— se evidencian limitaciones inherentes relacionadas con la iluminación, la presencia de ruido, y la variabilidad en los rostros y entonaciones, en su mayoría atribuibles a las especificaciones de la plataforma robótica Pepper.

6.1.4. Discusión

El desarrollo de este proyecto enfrentó múltiples desafíos técnicos y conceptuales que, aunque complejos, permitieron implementar soluciones creativas y establecer un camino claro para futuras mejoras. Uno de los obstáculos más significativos fue la incompatibilidad entre la plataforma robótica Pepper y las versiones actualizadas de ROS, lo que imposibilitaba la instalación directa del sistema operativo robótico. Para superar esta limitación, se diseñó una solución alternativa basada en un nodo ROS externo conectado mediante la dirección IP del robot, permitiendo así la ejecución remota del software. Sin embargo, esta aproximación introdujo una dependencia crítica de la estabilidad de la red y requirió ajustes continuos en la configuración para mantener la conexión activa.

En el ámbito del reconocimiento facial, se evaluaron inicialmente tres herramientas distintas, de las cuales las dos primeras, pese a su robustez, demandaban recursos computacionales excesivos para la capacidad de procesamiento de Pepper. La tercera opción, basada en un modelo YOLO entrenado con un dataset personalizado de 370 imágenes, demostró ser más adecuada al operar en un servidor externo, reduciendo así la carga sobre el robot. No obstante, el tamaño limitado del dataset afectó la precisión del sistema en condiciones no controladas, especialmente frente a expresiones faciales ambiguas o variaciones en la iluminación.

El reconocimiento de voz presentó sus propios desafíos, particularmente debido a la escasa documentación disponible sobre la API ALVoiceEmotionAnalysis. A pesar de esto, se lograron tasas de acierto del 81 % para alegría y 73 % para tristeza en entornos controlados. Sin embargo, el sistema mostró vulnerabilidad ante voces sintéticas o tonos neutros, generando falsos positivos.

Los resultados finales del proyecto validaron la viabilidad del modelo conceptual de empatía propuesto, el cual se estructura en cuatro fases claras: reconocimiento emocional, procesamiento de la información, generación de respuestas afectivas y cognitivas, y finalmente la interacción empática. La implementación de un sistema de votación ponderada, que asigna los pesos según el cuadro 4, demostró ser efectiva para integrar los distintos módulos.

Las pruebas realizadas en entornos controlados permitieron identificar áreas críticas para futuras mejoras. En primer lugar, la implementación de un filtro de audio robusto emerge como una necesidad prioritaria para reducir la sensibilidad del sistema al ruido ambiental. En segundo término, la ampliación del dataset de entrenamiento, actualmente limitado a 370 imágenes, resulta esencial para mejorar la precisión del reconocimiento facial en condiciones diversas. Asimismo, la inclusión de emociones adicionales más allá de alegría y tristeza, como ira o sorpresa, enriquecería significativamente el rango de interacciones posibles.

Mirando hacia el futuro, el proyecto podría beneficiarse sustancialmente de la incorporación de características avanzadas como el análisis de postura corporal mediante sensores 3D, que aportaría información valiosa sobre el lenguaje no verbal del usuario. Del mismo modo, la integración de modelos de procesamiento de lenguaje natural permitiría una comprensión más profunda del contexto conversacional, elevando la calidad de las respuestas generadas. Finalmente, la migración a ROS 2 podría resolver muchos de los problemas de conectividad actuales, ofreciendo mayor estabilidad y soporte para computación en tiempo real.

En conclusión, este proyecto no solo demostró la factibilidad de simular empatía en robots sociales mediante un enfoque multimodal, sino que también delineó un camino claro para su evolución futura. Las soluciones implementadas, como el nodo ROS remoto y el sistema de votación ponderada, sentaron bases sólidas para desarrollos posteriores. Sin embargo, la superación de limitaciones actuales, particularmente en el procesamiento de audio y la diversidad de datos, será crucial para alcanzar HRI verdaderamente naturales y efectivas, cumpliendo así con el objetivo último de ofrecer un soporte emocional genuino y adaptativo.

Capítulo 7

Conclusiones

De lo realizado podemos concluir

- Se logró establecer un modelo conceptual de empatía basado en cuatro fases clave (reconocimiento, proceso, resultado y respuesta empática), fundamentado en estudios psicológicos y adaptado a la interacción con robots. Este modelo permitió estructurar respuestas afectivas y cognitivas coherentes con las emociones detectadas (alegría y tristeza), utilizando técnicas validadas como la validación emocional. La integración de este marco teórico demostró ser esencial para simular comportamientos empáticos en el robot Pepper, mejorando la naturalidad de la interacción.
- Se desarrolló con éxito un software modular utilizando ROS que integró las herramientas de reconocimiento: facial (mediante redes neuronales entrenadas) y de voz (a través de la API ALVoiceEmotionAnalysis de Pepper). La arquitectura basada en nodos permitió superar limitaciones técnicas del hardware, como la baja capacidad de procesamiento del robot, al distribuir las tareas entre Pepper y un servidor externo. La votación ponderada asignó pesos específicos a cada canal optimizando la precisión en la detección emocional.
- Se validó el software mediante pruebas en entornos controlados, evidenciando un funcionamiento adecuado en condiciones ideales (precisión del 94 % en reconocimiento facial y 81 % en voz para alegría), aunque con limitaciones en escenarios reales debido a restricciones hardware de Pepper, como baja iluminación o ruido ambiental. Las métricas de desempeño (tiempos de inferencia <215 ms y confiabilidad >70 %) confirmaron la eficacia del modelo multimodal, respaldado por la votación ponderada para integrar resultados. La solución propuesta —ejecución remota vía

ROS— mitigó problemas de compatibilidad, pero destacó la necesidad de ampliar el dataset y mejorar filtros de audio para robustecer el sistema en futuras iteraciones.

Bibliografía

- [1] R. Camacho, *Comunicación empática: qué es y cómo puede mejorar tu vida*, 20 Mayo, 2019. dirección: <https://psicologiyamente.com/social/comunicacion-empatica>.
- [2] E. de Enciclopedia Significados, *Empatía*, 29 de Noviembre, 2023. dirección: <https://www.significados.com/empatia/>.
- [3] Á. B. Andrade, *¿Pueden los robots tener emociones?*, 15 de Febrero, 2022. dirección: <https://upadpsicologiacoaching.com/pueden-los-robots-tener-emociones/#:~:text=En%20conclusi%C3%B3n%2C%20sabemos%20que%2C%20a,que%20hacemos%20los%20seres%20vivos..>
- [4] M. A. Barnett, N. Eisenberg y J. Strayer, *Empathy and its development*, 1987.
- [5] A. I. Goldman, «Ethics and cognitive science», *Ethics*, vol. 103, n.º 2, págs. 337-360, 1993.
- [6] N. Eisenberg, N. D. Eggum, J. Decety y W. Ickes, «The social neuroscience of empathy», *Social neuroscience*, págs. 71-83, 2009.
- [7] N. F. Garzón Álvarez y J. C. Romero Moreno, «Implementación de un sistema de interacción humano-máquina que emplee lenguaje multimodal en una plataforma robótica basada en ROS.», Tesis doct., Universidad Santo Tomás.
- [8] H. Johanes, *Teoría del Conocimiento*, 1997.
- [9] J. S. Carvajal Amaya y F. O. Francis Stephenson, «Prototipo de un Chatbot web con inteligencia artificial de memoria limitada tipo II para interacción mediante lenguaje natural humano-máquina», 2020.
- [10] M. Saldarriaga Quintero, «Reconocimiento de sujetos y caracterización de sus emociones a partir de sus expresiones faciales usando algoritmos de reconocimiento facial», 2020.

-
- [11] Y. P. Esparza Rodríguez, «Implementación de un protocolo de intervención psicológica en pacientes pediátricos del Instituto del Corazón de Bucaramanga sometidos a cirugía cardíaca», 2013.
- [12] J. Martins, *Escucha activa: qué es y cómo practicarla con ejemplos*, Consultado: 2 de abril de 2025, feb. de 2025. dirección: <https://asana.com/es/resources/active-listening>.
- [13] E. R. Moraguez, *Robots Sociales Impresionantes: Interacción Humano-Robot en Entornos Reales*, Consultado: 2 de abril de 2025, 2025. dirección: <https://lovtechnology.com/robots-sociales-impresionantes-interaccion-humano-robot-en-entornos-reales/>.
- [14] D. L. A. Salgado e I. L. Gutiérrez, «Robótica social, la ciencia detrás de la interacción humana», *Revista +Ciencia*, n.º 28, págs. 19-21, mar. de 2022, Consultado: 2 de abril de 2025. dirección: <https://revistas.anahuac.mx/masciencia/article/view/1233>.
- [15] L. Barredo, I. Agyepong, G. Liu y S. Reddy, «Objetivo 3—Los Objetivos de Desarrollo Sostenible y un 2030 más saludable», *Crónica ONU*, abril de 2015, Consultado: 2 de abril de 2025. dirección: <https://www.un.org/es/chronicle/article/objetivo-3-los-objetivos-de-desarrollo-sostenible-y-un-2030-mas-saludable>.
- [16] O. de Desarrollo Sostenible Costa Rica, *Objetivo 3: Garantizar una vida sana y promover el bienestar para todos en todas las edades*, Consultado: 2 de abril de 2025. dirección: <http://ods.cr/objetivo/objetivo-3>.
- [17] N. del autor o autores, *Título del documento*, <https://repository.usta.edu.co/server/api/core/bitstreams/e7187ac4-cbd7-4f12-b185-c91c52a8aac1/content>, Accedido: 14 de abril de 2025, 2025.
- [18] A. P. Canazas, J. J. R. Blaz, P. D. T. Martínez y X. J. Mamani, «Sistema de identificación de emociones a través de reconocimiento facial utilizando inteligencia artificial», *Innovación y Software*, vol. 3, n.º 2, págs. 140-150, 2022.
- [19] M. León Nájera, «DISEÑO DEL ALGORITMO DEL PERFIL DEL ESTUDIANTE CON INTELIGENCIA ARTIFICIAL, INTEGRANDO EL RECONOCIMIENTO FACIAL»,
- [20] M. I. Gómez-León, «Desarrollo de la empatía a través de la Inteligencia Artificial Socioemocional», 2022.

-
- [21] O. Wahltinez, M. Rincón, M. Díaz-Mardomingo y S. García-Herranz, «Using deep learning to detect early signs of cognitive disease», en *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2020, págs. 3168-3173. DOI: 10.1109/SMC42975.2020.9283199.
 - [22] L. A. R. Franco, «Psicología cognitiva e inteligencia artificial: mitos y verdades», *Avances en Psicología*, vol. 22, n.º 1, págs. 21-27, 2014.
 - [23] G. M. Rickert, «Inteligencia artificial y psicología¿ Qué futuro nos espera?»,
 - [24] M. Demaría, S. Cusmai, A. Pérez Casas, S. Martins y M. Cavasotto, «Inteligencia artificial y psicología»,
 - [25] J. R. Altamirano-Yupanqui y A. Bernuy-Alva, «La inteligencia artificial en la Psicología: nuevos enfoques para la detección de las declaraciones falsas», *RISTI, Revista Ibérica de Sistemas e Tecnologías de Informação*, vol. 46, págs. 100-111, 2022.
 - [26] E. A. Huertas Tamayo, «Una relación conceptual entre filosofía e inteligencia artificial:¿ Deberían tener empatía los robots asistenciales?», 2020.
 - [27] J. Tao, T. Tan y R. W. Picard, eds., *Affective Computing and Intelligent Interaction: First International Conference, ACII 2005, Beijing, China, October 22-24, 2005, Proceedings* (Lecture Notes in Computer Science). Springer, 2005, vol. 3784, Consultado: 2 de abril de 2025, ISBN: 978-3-540-29621-8. DOI: 10.1007/11573548. dirección: <https://link.springer.com/book/10.1007/11573548>.
 - [28] Y. Cai, X. Li y J. Li, «Emotion Recognition Using Different Sensors, Emotion Models, Methods and Datasets: A Comprehensive Review», *Sensors*, vol. 23, n.º 5, pág. 2455, feb. de 2023, Consultado: 2 de abril de 2025. DOI: 10.3390/s23052455. dirección: <https://pubmed.ncbi.nlm.nih.gov/36904659/>.
 - [29] A. del documento, «Intelligent Facial Action and Emotion Recognition for Humanoid Robots», en *Proceedings of [Nombre de la Conferencia]*, Consultado: 2 de abril de 2025, Año de publicación, Número de páginas. DOI: 10.1109/[Identificadordelaconferencia].2014.6889647. dirección: <https://ieeexplore.ieee.org/document/6889647>.
 - [30] J. Huang, D. Yang, H. Wang y M. Xie, «Intelligent Facial Action and Emotion Recognition for Humanoid Robots», en *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Consultado: 2 de abril de 2025, 2015, págs. 2269-2274. DOI: 10.1109/ROBIO.2015.7419085. dirección: <https://ieeexplore.ieee.org/document/7280323>.

-
- [31] R. A. Brooks, «Elephants Don't Play Chess», Consultado: 2 de abril de 2025, Tesis doct., Massachusetts Institute of Technology, 1990. dirección: <https://dspace.mit.edu/handle/1721.1/46467>.
 - [32] A. del artículo, «Título del artículo», *Expert Systems with Applications*, vol. Volumen, n.º Número, Páginas, Año de publicación, Consultado: 2 de abril de 2025. DOI: 10.1016/j.eswa.2016.06.022. dirección: <https://www.sciencedirect.com/science/article/abs/pii/S0957417416304468>.
 - [33] J. Huang, D. Yang, H. Wang y M. Xie, «Intelligent Facial Action and Emotion Recognition for Humanoid Robots», en *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Consultado: 2 de abril de 2025, 2015, págs. 2269-2274. DOI: 10.1109/ROBIO.2015.7419085. dirección: <https://ieeexplore.ieee.org/document/8700376>.
 - [34] C. Breazeal, «Emotion and Sociable Humanoid Robots», *International Journal of Human-Computer Studies*, vol. 59, n.º 1-2, págs. 119-155, jul. de 2003, Consultado: 2 de abril de 2025. DOI: 10.1016/S1071-5819(03)00018-1. dirección: <https://www.sciencedirect.com/science/article/abs/pii/S1071581903000181>.
 - [35] H. Gunes y M. Piccardi, «Bi-modal emotion recognition from facial expression and body posture», en *Proceedings of the 3rd International Conference on Computer Vision Theory and Applications (VISAPP)*, Consultado: 2 de abril de 2025, 2008, págs. 500-507. dirección: <https://ieeexplore.ieee.org/document/5326306>.
 - [36] H. Gunes y M. Piccardi, «Affect Recognition from Face and Body: Early Fusion vs. Late Fusion», en *Human-Computer Interaction: New Developments*, N. Ayache, ed., Consultado: 2 de abril de 2025, IntechOpen, 2009, págs. 275-300. DOI: 10.5772/6439. dirección: <https://www.intechopen.com/chapters/186>.
 - [37] P. Ekman y W. V. Friesen, «Constants across cultures in the face and emotion», *Journal of Personality and Social Psychology*, vol. 17, n.º 2, págs. 124-129, 1971, Consultado: 2 de abril de 2025. DOI: 10.1037/h0030377. dirección: https://gunfreezone.net/wp-content/uploads/2008/12/Universal-Facial-Expressions-of-Emotions1-2015_03_12-21_10_38-UTC.pdf.
 - [38] P. Ekman y W. V. Friesen, *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*. Malor Books, 2003, Consultado: 2 de abril de 2025, ISBN: 978-1883536367. dirección: https://books.google.com/books/about/Unmasking_the_Face.html?id=TukNoJDgMTUC.

-
- [39] A. Lausen y K. Hammerschmidt, «Emotion recognition and confidence ratings predicted by vocal stimulus type and prosodic parameters», *Humanities and Social Sciences Communications*, vol. 7, n.º 2, 2020, Consultado: 2 de abril de 2025. DOI: 10.1057/s41599-020-0499-z. dirección: <https://www.nature.com/articles/s41599-020-0499-z>.
- [40] H. Fujisaki, *Prosody, Models, and Spontaneous Speech*. Springer, 1997, págs. 27-42, Consultado: 2 de abril de 2025. DOI: 10.1007/978-1-4612-2258-3_3. dirección: https://link.springer.com/chapter/10.1007/978-1-4612-2258-3_3.
- [41] D. A. Ochoa Muñoz y J. A. Pintado Yunga, «Desarrollo de un sistema prototipo de agente virtual para comunicación humano-máquina utilizando Bert», B.S. thesis, 2022.
- [42] M. I. Gómez-León, «Development of empathy through social emotional artificial intelligence», *Papeles del Psicólogo*, vol. 43, n.º 3, págs. 218-224, 2022.
- [43] Equipo de Asana, *Las 5 fases de la gestión de proyectos para que mejores el flujo de trabajo de tu equipo*, Consultado: 2 de abril de 2025, 2024. dirección: <https://asana.com/es/resources/project-management-phases>.
- [44] RecFaces, *Reconocimiento de emociones y expresiones faciales: Manejo Emocional*, Consultado: 2 de abril de 2025, 2024. dirección: <https://recfaces.com/es/articles/reconocimiento-de-emociones>.
- [45] J. Elías, *Análítica de emociones: cuando los robots comprendan tus sentimientos*, Consultado: 2 de abril de 2025, 2017. dirección: https://www.elespanol.com/omicrono/hardware/20170312/analitica-emociones-robots-comprendan-sentimientos/200230429_0.html.
- [46] Y.-H. Wang, J.-J. Lee, Y.-L. Chiu, C.-L. Lin, P.-C. Hsiao y C.-Y. Lin, «Validation of the Empathy Index in Taiwanese Clinical Nurses: A Cross-Sectional Study», *International Journal of Environmental Research and Public Health*, vol. 19, n.º 3, pág. 1889, 2022. DOI: 10.3390/ijerph19031889. dirección: <https://www.mdpi.com/1660-4601/19/3/1889>.
- [47] G. M. S. van Kleef, «Empathic Accuracy», en *International Encyclopedia of the Social & Behavioral Sciences*, J. D. Wright, ed., Elsevier, 2015, págs. 571-576. DOI: 10.1016/B978-0-08-097086-8.25074-0. dirección: <https://www.sciencedirect.com/science/article/abs/pii/B008043076701740X>.

-
- [48] J. Cheavens y otros, «The power of validation in helping people stay positive», *Journal of Positive Psychology*, 2020, Consultado: 2 de abril de 2025. dirección: <https://www.sciencedaily.com/releases/2020/12/201214123517.htm>.
 - [49] C. Bergland, «Want to Make Someone Feel Better? Validate Their Feelings», *Psychology Today*, 2020, Consultado: 2 de abril de 2025. dirección: <https://www.psychologytoday.com/us/blog/the-athletes-way/202012/want-to-make-someone-feel-better-validate-their-feelings>.
 - [50] A. Robotics, *Developer Center*, <https://www.aldebaran.com/developer-center/index.html>, Consultado: 2 de abril de 2025.
 - [51] C. Benitez, K. P. Howard y J. S. Cheavens, «The effect of validation and invalidation on positive and negative affective experiences», *The Journal of Positive Psychology*, vol. 17, n.º 1, págs. 46-58, 2022, Consultado: 2 de abril de 2025. DOI: 10.1080/17439760.2020.1832243. dirección: <https://www.tandfonline.com/doi/full/10.1080/17439760.2020.1832243>.
 - [52] C. E. Shenk y A. E. Fruzzetti, «The impact of validating and invalidating responses on emotional reactivity», *Journal of Social and Clinical Psychology*, vol. 30, n.º 2, págs. 163-183, 2011, Consultado: 2 de abril de 2025. DOI: 10.1521/jscp.2011.30.2.163. dirección: <https://guilfordjournals.com/doi/10.1521/jscp.2011.30.2.163>.
 - [53] M. M. Linehan, «Cognitive-Behavioral Treatment of Borderline Personality Disorder», *The Guilford Press*, 1993, Consultado: 2 de abril de 2025. dirección: <https://www.guilford.com/books/Cognitive-Behavioral-Treatment-of-Borderline-Personality-Disorder/Marsha-Linehan/9780898621839>.
 - [54] Y. Cai, X. Li y J. Li, «Emotion Recognition Using Different Sensors, Emotion Models, Methods and Datasets: A Comprehensive Review», *Sensors*, vol. 23, n.º 5, pág. 2455, 2023, Consultado: 2 de abril de 2025. DOI: 10.3390/s23052455. dirección: <https://pubmed.ncbi.nlm.nih.gov/36904659/>.
 - [55] G. López, J. A. Bagnoli, J. M. García, L. F. de la Ossa y V. Julián, «Multimodal emotion recognition with evolutionary computation for human-robot interaction», *Expert Systems with Applications*, vol. 64, págs. 585-598, 2016, Consultado: 2 de abril de 2025. DOI: 10.1016/j.eswa.2016.08.018. dirección: <https://www.sciencedirect.com/science/article/abs/pii/S0957417416304468>.
 - [56] P. Ekman, «Universal Facial Expressions of Emotion», *California Mental Health Research Digest*, vol. 8, n.º 4, Autumn de 1970.

-
- [57] A. S. Cowen y D. Keltner, «Semantic Space Theory: A Computational Approach to Understanding the Classification of Emotional Expressions», *Humanities and Social Sciences Communications*, vol. 7, n.º 1, págs. 1-11, 2020. DOI: 10.1057/s41599-020-0499-z. dirección: <https://www.nature.com/articles/s41599-020-0499-z>.
 - [58] I. H. Sarker, «Machine Learning: Algorithms, Real-World Applications and Research Directions», *SN Computer Science*, vol. 2, 2021, Consultado: 2 de abril de 2025. DOI: 10.1007/s42979-021-00592-x. dirección: <https://link.springer.com/article/10.1007/s42979-021-00592-x>.
 - [59] I. H. Sarker, «Machine Learning: Algorithms, Real-World Applications and Research Directions», *SN Computer Science*, vol. 2, 2021, Consultado: 2 de abril de 2025. DOI: 10.1007/s42979-021-00592-x. dirección: <https://link.springer.com/article/10.1007/s42979-021-00592-x>.
 - [60] I. H. Sarker, «Machine Learning: Algorithms, Real-World Applications and Research Directions», *SN Computer Science*, vol. 2, 2021, Consultado: 3 de abril de 2025. DOI: 10.1007/s42979-021-00592-x. dirección: <https://link.springer.com/article/10.1007/s42979-021-00592-x>.
 - [61] G. H. Yann LeCun Yoshua Bengio, «Deep learning», *Nature*, vol. 521, págs. 436-444, 2015. DOI: 10.1038/nature14539.
 - [62] AprendeIngenia, *Face Emotion Recognition*, Repositorio en GitHub, consultado: 3 de abril de 2025, 2025. dirección: <https://github.com/AprendeIngenia/face-emotion-recognition>.
 - [63] G. Sol, *Omes Tutorials - Reconocimiento de Emociones*, Repositorio en GitHub, consultado: 3 de abril de 2025, 2020. dirección: <https://github.com/GabySol/OmesTutorials2020/tree/master/7%20RECONOCIMIENTO%20DE%20EMOCIONES>.
 - [64] Roboflow, *Modelo de reconocimiento de Alegría y Tristeza*, Consultado: 3 de abril de 2025, 2025. dirección: <https://app.roboflow.com/juan-3mhxb/alegria-y-tristeza/models>.
 - [65] A. Robotics, *ALVoiceEmotionAnalysis - NAOqi Audio Module*, Consultado: 3 de abril de 2025, 2025. dirección: <http://doc.aldebaran.com/2-5/naoqi/audio/alvoiceemotionanalysis.html>.

-
- [66] A. del artículo, «Título del artículo», *Nature*, vol. Número de volumen si aplica, n.º Número de edición si aplica, Páginas si aplica, 2020, Consultado: 3 de abril de 2025. DOI: 10.1038/s41599-020-0499-z. dirección: <https://www.nature.com/articles/s41599-020-0499-z>.
- [67] J. D. A. Garzón, *Pepper - GitHub Repository*, <https://github.com/Reants/Pepper>, Repositorio consultado el 3 de abril de 2025, 2025.

Anexos de Implementación

.1. Conexión y control del robot Pepper

```
1 import qi
2 import rospy
3
4 session = qi.Session()
5 session.connect("tcp://192.168.0.102:9559")
6
7 speech = session.service("ALAnimatedSpeech")
8 photo = session.service("ALPhotoCapture")
9 memory_service = session.service("ALMemory")
10
11 def pituare():
12     photo.setResolution(2)
13     photo.setPictureFormat("jpg")
14     photo.takePicture("/home/nao/TesisEmociones", "IC")
15
16 def connect_and_speak():
17     rospy.init_node("Emotion_recognition")
18     rospy.loginfo("Iniciando nodo de emociones")
19     try:
20         print("Conectado correctamente")
21         speech.say("Conexi n exitosa, ahora puedo hablar contigo.")
22     except:
23         print("Error al conectar con Pepper")
24
25 if __name__ == "__main__":
26     connect_and_speak()
```

LISTING 1: Establecer conexión con plataforma robótica Pepper con ROS

```
1 import qi
2
3 # Crear una sesi n de conexi n
4 session = qi.Session()
5 # Configurar IP del robot
6 ip = "192.168.0.104"
7 # Conectarse al robot Pepper
8 session.connect("tcp://" + ip + ":9559")
```

LISTING 2: Establecer IP

```
1 import rospy
2 import time
3
4 # Inicializar nodo de ROS
5 rospy.init_node("Emotion_recognition")
6 # Registrar mensaje de inicio
7 rospy.loginfo("Nodo de ROS para Pepper iniciado.")
8
9 try:
10     # Registrar mensaje de conexi n
11     rospy.loginfo("Conexi n a Pepper establecida.")
12     time.sleep(3)
13     # Inicializar variable
14     saludo_recibido = False
15 except Exception as e:
16     rospy.logerr("Error al conectar con Pepper: {}".format(e))
```

LISTING 3: Conexi3n con Pepper completa

.2. Reconocimiento de emociones por voz con Pepper

```
1 #!/usr/bin/env python
2 # -- encoding: UTF-8 --
3 import qi
4 import time
5
6 # Establece la sesi n y conecta con el robot
7 session = qi.Session()
8 session.connect("tcp://127.0.0.1:9559")
9
10 # Inicializa los servicios
11 emotion_analysis_service = session.service("ALVoiceEmotionAnalysis")
12 memory_service = session.service("ALMemory")
13 animated_speech_service = session.service("ALAnimatedSpeech")
14
15 # Suscribe al an lisis de emoci n en la voz
16 emotion_analysis_service.subscribe("EmocionDeVoz")
```

LISTING 4: Análisis de emoción por voz con Pepper - parte 1

```

1 try:
2     while True:
3         print("Escuchando emoci n en la voz...")
4         # Obt n el evento EmotionRecognized desde ALMemory
5         emotion_data = memory_service.getData("ALVoiceEmotionAnalysis/
EmotionRecognized")
6         if emotion_data:
7             matched_emotion_index = emotion_data[0][0] # ndice de la
emoci n dominante
8             matched_emotion_level = emotion_data[0][1] # Nivel de la
emoci n dominante
9
10            # Imprime la emoci n detectada y el nivel
11            print("Emoci n detectada ( ndice ): {}, nivel: {}".format(
12                matched_emotion_index, matched_emotion_level))
13
14            # Responde en funci n de la emoci n dominante
15            if matched_emotion_index == 3: # Joy
16                animated_speech_service.say("Parece que est s muy
alegre, me alegra escucharlo!")
17            elif matched_emotion_index == 4: # Sorrow
18                animated_speech_service.say("Lamento que est s triste.
Estoy aqu para animarte.")
19
20            # Espera un poco antes de analizar nuevamente
21            time.sleep(3)
22 except KeyboardInterrupt:
23     # Desuscribe en caso de interrupci n (CTRL+C)
24     emotion_analysis_service.unsubscribe("EmocionDeVoz")
25     print("Suscripci n finalizada.")

```

LISTING 5: Análisis de emoción por voz con Pepper - parte 2

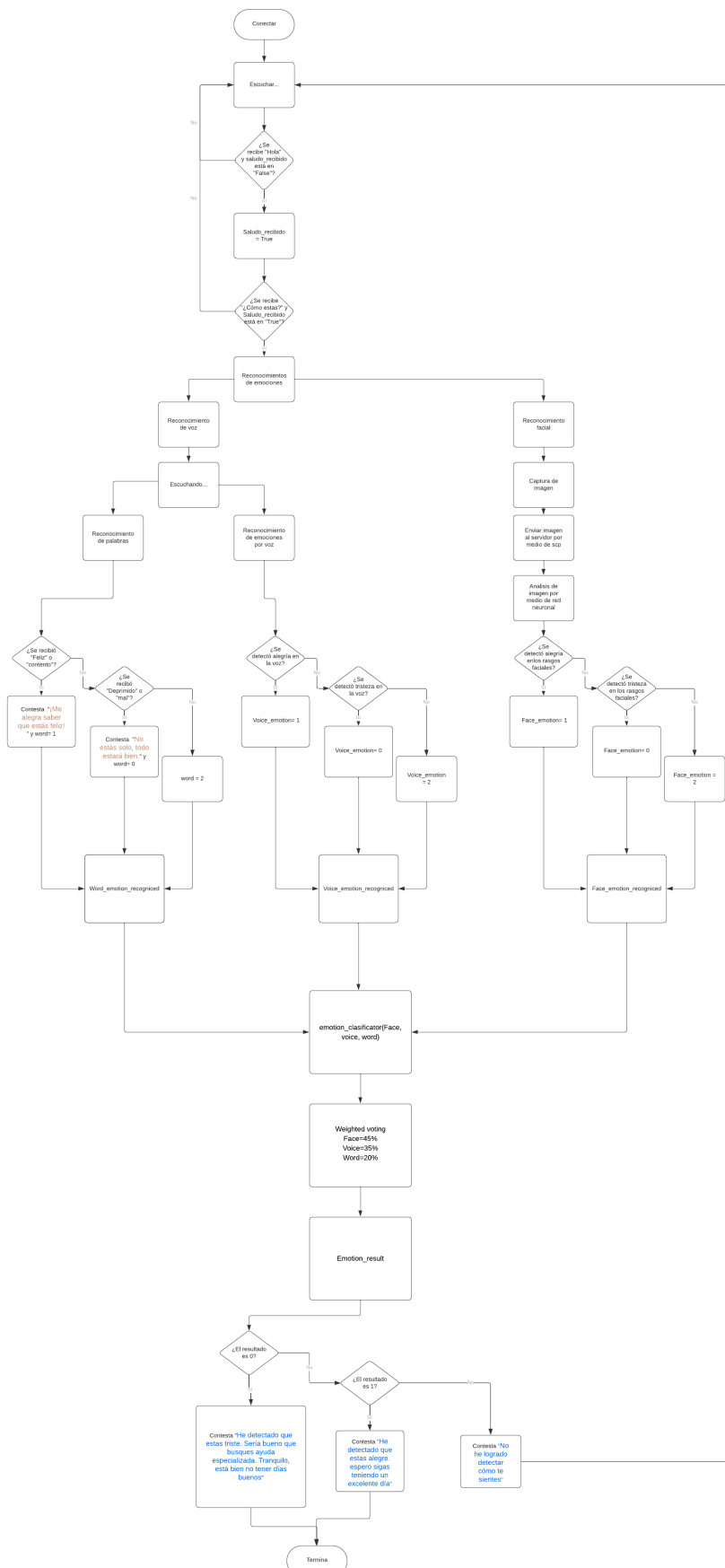


FIGURA 25: Diagrama de flujo del sistema