

DECIMOQUINTA CONVOCATORIA PARA EL FOMENTO DE LA INVESTIGACIÓN Y LA INNOVACIÓN 2020

Título del proyecto	
Herramienta para Reconocimiento de Objetos para la Plataforma Robótica Pepper	
Campo de acción	Transdisciplinariedad - Aporte al PIM
Desarrollo tecnológico con apuesta social	<p>Línea 3: Proyección social e investigación pertinentes. Los productos del proyecto se orientan a la generación de nuevo conocimiento y su objetivo general apunta hacia el mejoramiento de la calidad de vida de la sociedad en general. En específico se espera contribuir con el Campo de <u>Acción</u> de la Universidad “Sociedad”:</p> <ul style="list-style-type: none"> ▪ Garantizar que los programas y proyectos de proyección social se articulen con las demás funciones sustantivas. ▪ Incrementar la producción investigativa con impacto regional, nacional e internacional. ▪ Aumentar de manera progresiva y sostenible la cooperación académica con otras instituciones nacionales e internacionales.
Articulación con funciones sustantivas y el sector social y productivo	
<p>Aporte a la docencia: Apropriación de conocimientos para aplicarlos a los contenidos curriculares y espacios académicos. En los espacios académicos Visión Artificial, Introducción a la Inteligencia Artificial y Robótica, ya que la investigación está enfocada directamente en estas áreas del conocimiento, es así que el desarrollo de herramientas software para Reconocimiento de Objetos contribuirá a mejorar los temarios y prácticas de laboratorio de los espacios académicos.</p> <p>Aportes a la investigación: El alcance de proyecto es adecuado para que a partir de este se generen sub-proyectos en los cuales se vinculen estudiantes de pregrado en proyecto de grado, proyecto de semillero y/o proyecto de joven investigador.</p> <p>Este proyecto hace parte del macro proyecto de Robótica Social del Grupo de Estudio y Desarrollo en Robótica – GED, el cual se encuentra en su segunda fase de desarrollo, es así que se espera los resultados obtenidos permitan formular nuevos proyectos de investigación, para financiamiento interno y apoyados por el Convenio SINFONIA, del cual hace parte la Universidad.</p> <p>Además, el proyecto genera un desarrollo tecnológico que servirá para nuevas investigaciones. Con los productos esperados se contribuirá al fortalecimiento del grupo de investigación GED.</p> <p>Aporte a la responsabilidad social universitaria: El proyecto se enmarca dentro del Convenio SINFONIA, integrado por Bancolombia y las Universidades de los Andes, del Magdalena y Santo Tomás. El desarrollo del proyecto contribuirá al cumplimiento del objeto del Convenio así aportando a la articulación con el sector productivo.</p>	

Además, el proyecto apunta hacia el mejoramiento de la calidad de vida de la sociedad a través de Robótica Social, se espera generar actividades de: participación ciudadana, apropiación social del conocimiento, promoción y divulgación en el ámbito nacional e internacional, como: conferencias, simposios, talleres, notas de prensa, entre otras actividades.

Grupo de investigación	Línea de investigación en la que se inscribe el proyecto
Grupo de Estudio y Desarrollo en Robótica – GED	Visión Artificial

Nombre del Investigador principal	Enlace CvLAC	Enlace ORCID	Enlace Google Académico
Fabian Eduardo Pérez Gordillo	http://scienti.colciencias.gov.co:8081/cvlac/visualizad/or/generarCurriculoCv.do?cod_rh=0001516111	https://orcid.org/0000-0002-2746-8733	https://scholar.google.com/citations?user=vncSAb0AAAAJ&hl=es
División	Facultad	Programa	Grupo de investigación
Ingenierías	Ingenierías Electrónica	Ingeniería Electrónica	GED
Nombre del Co-investigador	Enlace CvLAC	Enlace ORCID	Enlace Google Académico
Edgar Camilo Camacho Poveda	http://scienti.colciencias.gov.co:8081/cvlac/visualizad/or/generarCurriculoCv.do?cod_rh=0001630084	https://orcid.org/0000-0002-6084-2512	https://scholar.google.com/citations?user=tJG988kAAAAJ&hl=es
División	Facultad	Programa	Grupo de investigación
Ingenierías	Ingenierías Electrónica	Ingeniería Electrónica	GED

Resumen de la propuesta	Palabras clave
<p>El Grupo de Estudio y Desarrollo en Robótica – GED de la Facultad de Ingeniería Electrónica se encuentra desarrollando un macro-proyecto en Robótica Social utilizando la plataforma robótica Pepper como un robot de servicio, en ambientes cotidianos del hombre y la mujer, para que este que cuente con capacidades básicas para desarrollar tareas comunes que permitan mejorar el bienestar de los seres humanos.</p> <p>Enmarcado en este objetivo esta propuesta plantea desarrollar un grupo de herramientas software útiles para la ejecución de tareas de la plataforma robótica Pepper que incluyan: herramientas para reconocimiento de objetos, herramientas para la construcción del DataSet, la definición de categorías y clases para entrenamiento y que, además integre funcionalidades adicionales para proporcionar información específica de los objetos.</p>	Clasificación y Reconocimiento de objetos, DataSet de Entrenamiento, Robots de Servicio, Aprendizaje Supervisado.

Problema de investigación

Los robots de servicio o sociales realizan tareas útiles para el bienestar de los seres humanos en escenarios cotidianos como: el hogar, centros comerciales, restaurantes y universidades, entre otros. En el cumplimiento de este objetivo se reúnen varias disciplinas que van desde el diseño mecánico, la inteligencia artificial, la visión artificial hasta la psicología, entre otras.

En la ejecución de la mayor parte de las tareas de un robot de servicio, la visión artificial es una de las herramientas más empleadas. Teniendo en cuenta que los robots de servicio se emplean en escenarios cotidianos para el hombre, contar con una herramienta de visión artificial que tenga la capacidad de reconocer cualquier objeto en cualquiera de estos escenarios es actualmente una tarea demasiado compleja, por la gran variedad de objetos que se pueden encontrar en cada uno de estos escenarios. Por tal motivo, actualmente los objetos que un robot de servicio puede reconocer están limitados a los posibles objetos que se encuentren en el escenario donde el robot ejecuta sus tareas. Por ejemplo, un robot de servicio doméstico cuyo ambiente es un hogar, contará con la capacidad de reconocer objetos típicos de este ambiente como: televisores, camas, entre otros; sin embargo, normalmente no podrá reconocer un tablero o pupitre o cualquier otro objeto atípico de un hogar. Es por ello que cada vez que un robot de servicio cambia de ambiente es necesario “enseñarle” a reconocer los nuevos objetos que podría encontrar en el nuevo escenario [1].

Para el reconocimiento de objetos se emplean herramientas de aprendizaje supervisado, las cuales necesitan de bases de datos – *DataSet* de gran tamaño para el entrenamiento adecuado, es más, el reconocimiento de objetos está fuertemente influenciados por la calidad del *DataSet* [1]. La construcción de estos *DataSet* se realiza generalmente tomando imágenes de la Web o con fotos para ser etiquetadas después, en efecto, no es una tarea trivial, que puede ser tediosa y tomar bastante tiempo.

Sin dejar de lado que para algunas tareas no es suficiente con únicamente reconocer el objeto, algunas de ellas requieren información específica del objeto, por ejemplo, para realizar una tarea que involucre manipulación del objeto, es necesario proporcionar la ubicación espacial de dicho objeto.

Este proyecto entonces se centra específicamente en visión artificial, con el fin de desarrollar y/o integrar un grupo de herramientas software útiles para la plataforma robótica Pepper que incluyan, pero sin limitarse a herramientas para: i) reconocimiento de objetos, ii) construcción de *DataSets*, y que además iii) integre funcionalidades adicionales para proporcionar información específica de los objetos.

Teniendo en cuenta lo argumentado anteriormente, nos lleva a plantearnos la siguiente pregunta de investigación:

¿Cómo implementar un conjunto de herramientas de visión artificial para el reconocimiento de objetos para un para la plataforma robótica Pepper?

Justificación

Pepper es un robot social humanoide desarrollado por SoftBank Robotics, optimizado para Interacción Humano-Robot – HRI [2]. La Universidad Santo Tomás como miembro del Convenio SINFONIA integrado también por Bancolombia y las Universidades del Magdalena y de los Andes, cuenta con uno de estos robots (ver ANEXO 1). Con esta plataforma, el Grupo de Estudio y Desarrollo en Robótica – GED se encuentra un macro proyecto en robótica social enmarcado dentro de los objetivos de SINFONIA, el cual tiene como meta emplear la plataforma robótica Pepper como un robot de servicio, en ambientes cotidianos del hombre y la mujer, para que este que cuente con capacidades básicas para desarrollar tareas comunes que permitan mejorar el bienestar de los seres humanos.

Un adecuado conjunto de herramientas de visión artificial permitiría: i) contar con las capacidades adecuadas para el reconocimiento y clasificación de objetos, ii) mitigar las dificultades técnicas y logísticas que se presentan en la construcción de *DataSets*, y además iii) integrar funcionalidades adicionales como: proporcionar información específica de los objetos, por ejemplo: color, ubicación espacial que son necesarias para manipulación, localización y demás

tareas del robot. Brindando así un conjunto de herramientas útiles y efectivas para el reconocimiento de objetos que necesita la plataforma robótica Pepper para el correcto desarrollo de las tareas en diferentes ambientes.

En términos investigativos la propuesta se soporta en lineamientos internacionales para robótica como los presentados en [3] y [4] cuyas directrices o retos en robótica apuntan a la integración de los robots en la cotidianidad de la sociedad, la interacción natural y la cooperación humano-robot. Además, de contar con el soporte, metodología y cooperación académica de SINFONIA y la red mundial de investigación *RoboCup* [5], en específico de la liga *Robocup@Home* cuya competencia de robots de servicio es la más importante a nivel mundial y de la cual ya se ha tenido la experiencia de participar en 2019 con el equipo Sinfonia Pepper Team [6] [7].

Objetivo general

Implementar un conjunto de herramientas de visión artificial para la plataforma Robótica Pepper, que integre el reconocimiento y clasificación de objetos con funcionalidades adicionales para el entrenamiento, creación de bases de datos de imágenes y que tenga capacidad de brindar información característica de los objetos; empleando clasificadores basados en técnicas de aprendizaje máquina y métodos de procesamiento digital de imágenes.

Objetivos específicos

- Seleccionar una herramienta para la clasificación y reconocimiento de objetos ya implementada basada en técnicas de aprendizaje supervisado que permita el entrenamiento de nuevas categorías y clases de objetos.
- Implementar una herramienta software para la construcción de bases de datos de imágenes para el entrenamiento de un clasificador de objetos a partir de fotografías, videos o imágenes.
- Implementar una herramienta de visión artificial que permitan brindar información de las características específicas de los objetos reconocidos por el clasificador.
- Integrar las herramientas de clasificación y reconocimiento de objetos, construcción de bases de datos de imágenes para el entrenamiento, y de visión artificial para brindar información de las características específicas de los objetos; en una sola herramienta para la plataforma robótica Pepper.
- Realizar pruebas de la herramienta de clasificación y reconocimiento de objetos con la plataforma robótica Pepper en un ambiente con objetos reales y evaluar cuantitativamente los resultados.

Estado del arte y marco conceptual

ROBOTS DE SERVICIO

La robótica de asistencia o de servicio no ha sido definida formalmente aún, pero un robot de servicio es aquel que realiza varias y diferentes tareas de manera autónoma interactuando de manera natural con: las personas, los objetos y ambientes reales [8], [9], [10]. Los ambientes pueden ser, conocidos, o parcialmente conocidos incluso en escenarios desconocidos totalmente.

Un enfoque y una de las principales dificultades, es la interacción natural con el mundo real, a diferencia de la robótica industrial en la que el entorno y las condiciones son conocidas, los robots de servicio deben hacer frente a una gran cantidad de incertidumbre. Además, los objetos y las personas pueden moverse, desaparecer y reaparecer, es decir, el entorno es dinámico. El robot debe ser capaz de manipular objetos en diferentes localizaciones, de desplazarse en diferentes terrenos. Al interactuar con las personas debe poseer cierta inteligencia básica (social) y contar con la capacidad de distinguir diferentes

personas. Por último, pero no menos importante, el funcionamiento seguro y robusto de estos sistemas en entornos tan inciertos y dinámicos es un requisito fundamental para su futura aceptación y aplicabilidad general [10].

La construcción de estos robots autónomos requiere de la integración de un gran conjunto de herramientas y tecnologías entre ellas: Interacción Humano-Robot – HRI (reconocimiento del habla, reconocimiento de gestos, entre otros), navegación y mapeo, razonamiento, planificación, control de comportamiento, reconocimiento de objetos, manipulación de objetos [8], [10]. Con respecto a la inteligencia artificial, los sistemas deben contener comportamientos adaptativos pero sólidos y métodos de planificación, inteligencia social y capacidades de aprendizaje.

Aunque son muchos los campos del conocimiento necesarios para implementar un robot de servicio, recientemente se ha realizado un progreso en cada una de estas áreas en términos de hardware y software como los presentados en [11], [12], [13] [14]. En muchas ocasiones potenciados por comunidades de investigación internacionales como RoboCup@Home. Pero el reto en robótica social no consiste entonces en la integración y reutilización de tecnología existente, pero puede ayudar a optimizar tiempo y esfuerzo a los investigadores, quienes pueden centrarse en un campo de investigación particular mientras mantienen una plataforma robótica completamente operable [10]. Ejemplos prominentes de plataformas robóticas estandarizadas o propietarias para asistencia y servicio incluyen los proyectos como: ReadyBot [15], PR2 [16], Wakamaru [17], PaPeRo [18], Toyota HSR [19] y Pepper [20].

RECONOCIMIENTO DE OBJETOS

El reconocimiento de objetos ha sido desde su inicio y actualmente uno de los principales y más estudiados tópicos en visión artificial, sus aplicaciones son muchas y variadas que van desde indexación y recuperación de imágenes, a video-vigilancia, robótica y medicina. Se podría decir que reconocimiento de objetos consiste en detectar en una imagen un objeto y luego asignarle una determinada clase a partir de una agregación de características que un clasificador ha sido entrenado para reconocer [21]. La capacidad de reconocimiento de objetos está presente dentro las funcionalidades necesarias con las debe contar el robot para llevar a cabo muchas de las tareas que ejecuta.

La comunidad de visión artificial se ha dedicado mucho esfuerzo en diferentes enfoques para diseñar poderosas herramientas para implementar clasificadores efectivos con altas precisiones [21]. la mayor parte de estos enfoques se basan en técnicas de aprendizaje de máquina especialmente en las técnicas de aprendizaje por refuerzo, dentro de las que destacan: K-ésimo Vecino más Cercano – KNN, Máquina de Vector de Soporte – SVM o Redes Neuronales Convolucionales – CNN, esta última popularizada por su alto desempeño ya que ha mostrado ser excelente como extractor de características [21].

En concordancia con el desarrollo de reconocimiento de objetos, actualmente se encuentran disponibles varias herramientas para clasificación y reconocimiento de objetos con la CNN ya implementada en software, que además cuenta con las herramientas necesarias para realizar el entrenamiento, y definición de categorías y clases de objetos, algunas de ellas ya previamente entrenadas, entre las que destacan: AlexNet [22], GoogleNet [23], YOLO [24], SSD [25], RetinaNet [26]. En [26] y [27] se presentan una comparación cuantitativa de algunas de estas herramientas empleando la compensación de velocidad/precisión de acuerdo a las métricas mAP de .5 IOU.

Conjunto de Imágenes de Entrenamiento

Tradicionalmente los clasificadores son a menudo entrenados con conjuntos de imágenes que son construidos específicamente por expertos en visión artificial cuya experiencia les orienta en la elección de las imágenes para que cumplan con ciertas características (instancias prototípicas de clase). Por supuesto tal elección de imágenes no es arbitraria, pero los criterios de elección se dejan implícitos que también son los criterios de identidad del objeto que se desea detectar [21]. En este tipo de metodología para la construcción del DataSet se presenta un inconveniente relevante ya que la cantidad de imágenes necesarias para el entrenamiento es grande, lo que lo hace un proceso costoso, además, en ocasiones según el tipo de objeto a detectar el número de imágenes que se consigue recopilar es limitado.

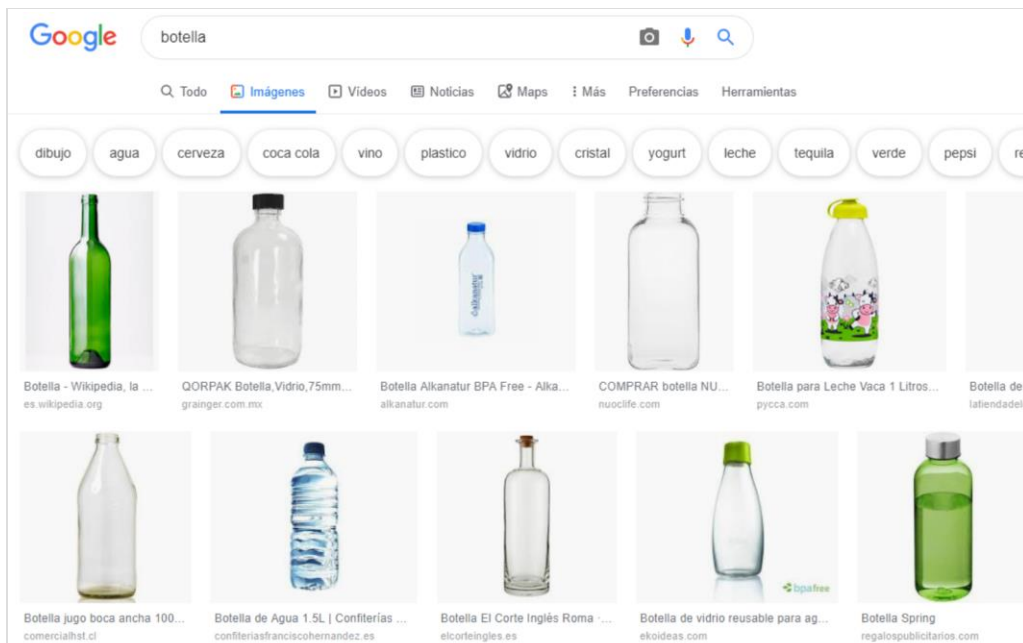


Figura 1. Imágenes obtenidas en Google Imágenes con la búsqueda de “botella”.

Para mitigar la desventaja de tener pocas imágenes de entrenamiento por clase y, en general, pocas clases de objetos, en los últimos años han surgido proyectos que explotan la llamada "wisdom of crowd" para poblar conjuntos de imágenes de entrenamiento, a través de métodos de recolección de imágenes descargadas de la Web [21], [28], [29].

Por otra parte, uno de los más importantes proyectos para construcción de DataSet de entrenamiento basados en la Web es “ImageNet” [30]. Actualmente ImageNet es el DataSet de imágenes más grande disponible para la comunidad de visión artificial [21].

En términos muy generales la idea de ImageNet y otros proyectos como el presentado en [31], consiste en utilizar un término que describa una clase de un objeto que se desea reconocer, por ejemplo botella, plato, vaso, entre otros y utilizarlo como palabra clave en el motor de búsqueda de imágenes (Google) en Internet y con las imágenes recuperadas construir el conjunto de imágenes de la clase para entrenamiento. En muchos casos este enfoque no es el más adecuado ya que se obtienen imágenes heterogéneas muchas veces sin un escenario de fondo lo que las hace instancias no prototípicas ver Figura 1.

SISTEMA OPERATIVO PARA ROBOTS – ROS

ROS (Robot Operating System) provee librerías y herramientas para ayudar a los desarrolladores de software a crear aplicaciones para robots. ROS provee abstracción de hardware, controladores de dispositivos, librerías, herramientas de visualización, comunicación por mensajes, administración de paquetes y más. ROS está bajo la licencia open source, BSD. ROS actualmente corre sobre plataformas basadas en Unix, es principalmente probado y con versiones estables en sistemas Ubuntu y Mac OS [32].

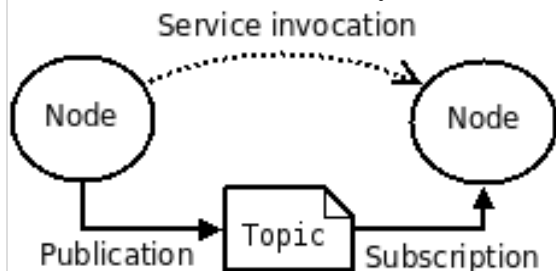


Figura 2. Ejemplo de Operación ROS.

La ejecución de ROS funciona como una red par-a-par de procesos lo que le permite procesamiento distribuido, para ello hace uso de diferentes métodos de comunicación incluyendo comunicación RPC, transmisión asíncrona de datos. Es importante tener presente que ROS no es un marco de ejecución de tiempo-real, aunque es posible integrar funcionalidades de tiempo-real [32].

Como sistema operativo conceptualmente ROS maneja 3 niveles: el nivel de Archivos de Sistema, nivel Computacional Gráfico y nivel de Comunidad. El primero de los conceptos los Archivos de Sistema que principalmente cubre los recursos encontrados en disco, tales como: Paquetes, Meta-paquetes, Manifiesto, Repositorios, Mensajes y Servicios [32].

El nivel Computacional Gráfico es manejado como una red par-a-para el procesamiento de datos realizado a través de conceptos propios de ROS como: Nodos, Maestro, Servidor de Parámetros, Mensajes, Tópicos, Servicios y Maletas. El Maestro de ROS actúa como un nombre-de-servicio, este almacena la información de registro de Tópicos y Servicios para los Nodos ROS. Los Nodos comunican al Maestro reportes su información de registro, también reciben del Maestro información acerca de otros nodos y conexiones según sea apropiado. El Maestro puede hacer llamadas a los nodos para registro de nuevos cambios, lo cual permite crear dinámicamente conexiones [32].

Los Nodos se conectan directamente a otros Nodos, el Maestro únicamente proporciona información de búsqueda como un servidor DNS. Los Nodos que se suscriben a un tópico solicitarán las conexiones a los Nodos que publican en ese tópico, y la conexión se establecerá empleando el protocolo TCP/IP [32].

Adicionalmente ROS cuenta con desarrollo de diferentes librerías para dar soporte a diferentes tareas y herramientas comunes para robots, entre ellas se encuentran librerías para visión artificial, una de las más empleadas es Vision_OpenCV para ROS [33], la cual proporciona el empaquetado de la popular biblioteca OpenCV para ROS.

PLATAFORMA ROBÓTICA PEPPER

Pepper es una plataforma robótica la compañía japonesa SoftBank Robotics de 120 cm de alto, optimizado para la Interacción Humano-Robot provocando un sentimiento de empatía con el operador humano, suponiendo un gran paso evolutivo hacia la inserción de los robots en entornos comerciales. Es capaz y es capaz de interactuar con las personas a través de la conversación y su pantalla táctil [2]. Sus funcionalidades le brindan la capacidad de interactuar de manera natural con los humanos, es decir, emplear comunicación multimodal, interpretar el lenguaje verbal como el no verbal.

La interactividad es el concepto principal en el diseño de Pepper. Sus múltiples sensores táctiles y de sonido, junto a sus cámaras RGB y de profundidad, le permiten registrar e interpretar de forma detallada su ambiente. Además, su pantalla táctil transmite y recibe información y posee conexión on-line. Este potente hardware ofrece múltiples opciones a la hora de programar y configurar el robot para que interactúe con su entorno.

El robot puede ser programado bajo diferentes lenguajes de programación (Choreographe, Python y C++), que lo hacen óptimo como plataforma robótica al mismo tiempo que lo convierten en una herramienta perfecta para investigación y desarrollo de nuevas aplicaciones.

Además, potenciado por RoboCup@Home muchos investigadores han desarrollado varias y variadas aplicaciones y casos de uso con este robot, lo que brinda un soporte para la investigación en robótica social con el uso de esta plataforma robótica [8].

Metodología

Considerando el objetivo del proyecto se pretende hacer uso de varias fases investigativas, de manera tal que, se pueda alcanzar el cumplimiento de los objetivos propuestos, así se plantean las siguientes Fases:

Fase 1: Recolección de Información

La etapa inicial del proyecto se centrará en la recolección y revisión de información acerca de visión artificial, en específico la relacionada con reconocimiento de objetos. Se revisará las diferentes herramientas de aprendizaje de máquinas empleadas para el reconocimiento y clasificación de objetos, sin embargo, esta revisión de información se focalizará en Redes Neuronales Convolucionales – CNNs, ya que esta técnica de aprendizaje supervisado es la más utilizada para reconocer y clasificar objetos [1] [21] [34]. También, se hará una búsqueda y revisión de las técnicas empleadas para la construcción de las bases de datos *DataSet* de imágenes que son utilizadas para el entrenamiento haciendo énfasis en las características y requerimientos recomendados. Es importante resaltar que durante esta fase aparte de documentación se buscará: clasificadores de objetos ya implementados sin entrenamiento, pre-entrenadas, o entrenadas completamente. Estas últimas, serán útiles en etapas posteriores para realizar actividades de prueba y evaluación. También se buscarán *DataSet* de imágenes

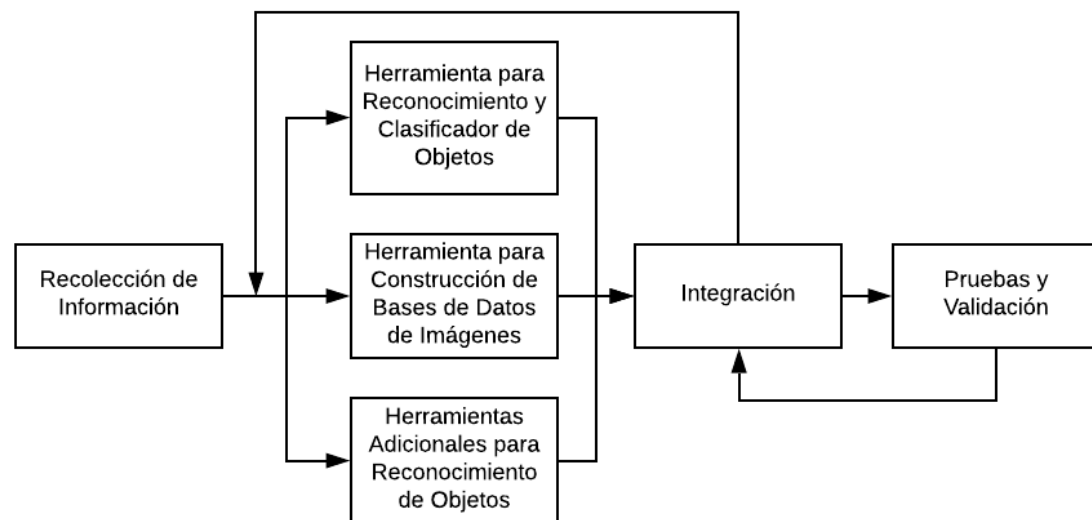


Figura 3. Metodología para ejecución del Proyecto.

de objetos de ambientes comunes para los robots de servicio ya construidas que puedan ser modificadas para en etapas posteriores realizar evaluación de la herramienta para construcción del *DataSet* y también para realizar pruebas de la totalidad del conjunto de herramientas para reconocimiento de objetos.

Fase 2: Selección de Herramienta para Reconocimiento y Clasificador de Objetos

Con la revisión del estado del arte, se contará con las herramientas necesarias para la selección de una herramienta ya implementada para reconocer y clasificar objetos que permita el entrenamiento de nuevas categorías y clases de objetos, pero que podría ser o no pre-entrenada. Inicialmente se plantea el uso de YOLO [24] [27] que aparte del clasificador ya implementado cuenta con las herramientas necesarias para el entrenamiento.

Según las necesidades encontradas durante la ejecución del proyecto, podrá evaluarse el uso de otra herramienta para clasificación y reconocimiento de objetos que cuente con características similares a YOLO. En esta fase se realizarán las primeras actividades de entrenamiento del clasificador seleccionado utilizando *DataSet* ya contruidos.

Fase 3: Herramienta para Construcción de Bases de Datos de Imágenes

Basado principalmente en los resultados de la primera fase respecto a los *DataSet* de imágenes se iniciará con el desarrollo y/o integración de herramientas software que permita la construcción de *DataSets* de imágenes a partir de fotografías o videos tomados por nosotros mismos o por imágenes descargadas de la Web. La herramienta desarrollada contará con funcionalidades para: descarga de imágenes de la Web, para el etiquetado de imágenes, y para la construcción de imágenes a partir de un archivo de video. Es importante resaltar que, en esta fase los *DataSets* contruidos se evaluarán para que cumplan con las características y requerimientos necesarios (número de imágenes por clase, tamaño de la imagen, resolución, escenario de la imagen, entre otros) para entrenar el clasificador seleccionado en la etapa anterior.

Fase 4: Herramientas Adicionales para Reconocimiento de Objetos

En esta etapa del proyecto se desarrollarán y/o integrarán algoritmos de visión artificial que permitan brindar información adicional de las características de los objetos, para complementar el conjunto de herramientas de reconocimiento de objetos. Estas herramientas pueden incluir, pero sin limitarse a: algoritmos de enmascaramiento para la detección de color, algoritmos para reconocimiento de texto y algoritmos para determinar la ubicación espacial, entre otros. La integración de estos algoritmos constituirá una única herramienta que entregará la información adicional de las características específicas de cada uno de los objetos que el robot esté reconociendo en cierto instante.

Fase 5: Integración

Esta fase iniciará en el momento en que se tengan las primeras versiones de las herramientas de las fases 2, 3 y 4. Esta fase la constituyen, actividades de integración de herramientas software basados en la arquitectura que se presenta en la Figura 4, para constituir una sola herramienta con todas las funcionalidades de las demás y que además cuente con una documentación adecuada para el usuario final, para ser utilizada a través de comandos o por interfaz de usuario.

Como la herramienta está orientada a ser utilizada por la plataforma robótica Pepper, la integración se realizará a través de nodos de ROS en el que cada nodo constituye una o varias herramientas y que en conjunto todos los nodos constituyen un paquete ROS que será la herramienta para el reconocimiento de objetos. Es importante mencionar que adicional a estos nodos, se desarrollaran uno o más nodos de ROS en la plataforma robótica Pepper para obtener los datos de las cámaras RGB y de profundidad del robot.

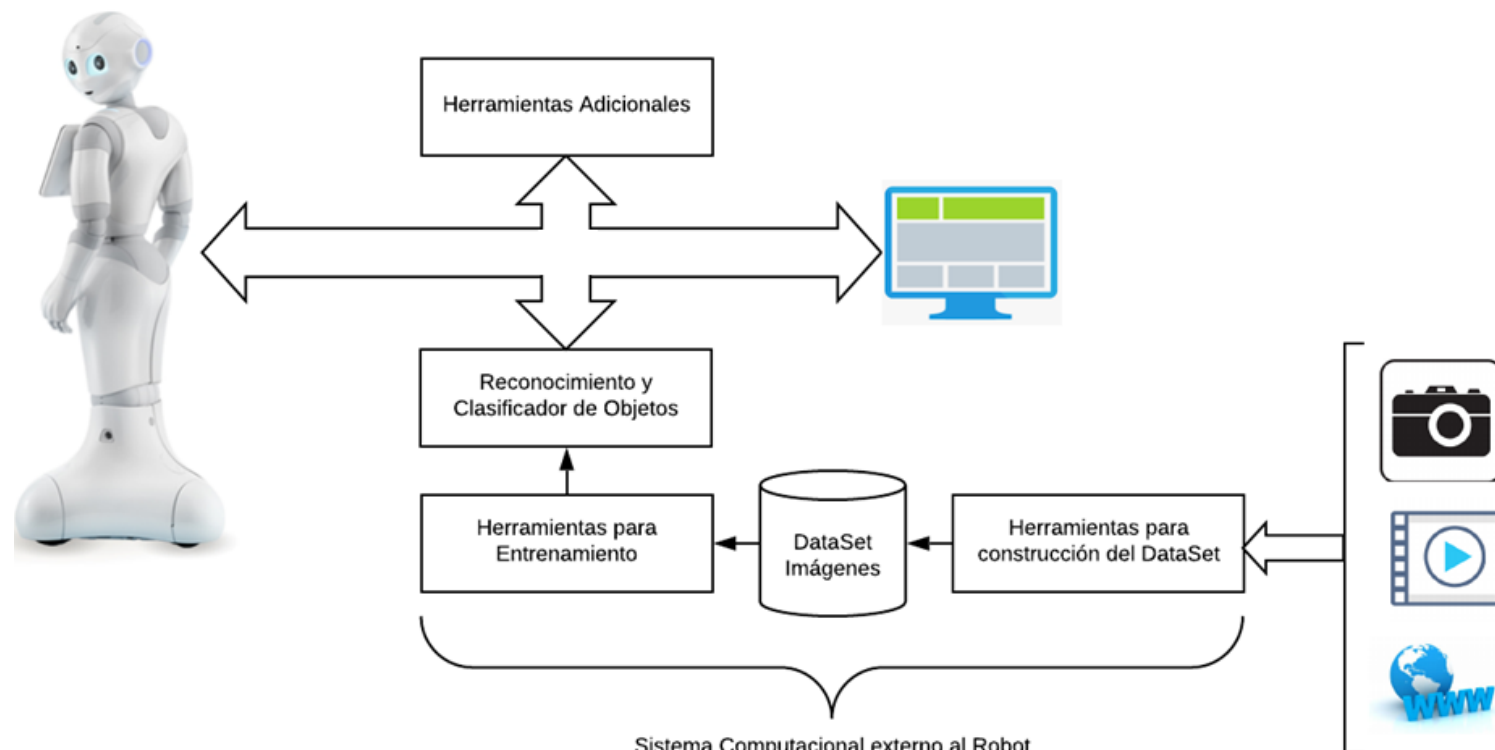


Figura 4. Arquitectura de la Solución.

Fase 6 Pruebas y Validación

El objetivo de esta fase final es ejecutar el paquete de ROS para reconocimiento de objetos utilizando la plataforma robótica Pepper, se realizarán pruebas empleando objetos reales de un ambiente cotidiano para un robot de servicio. Los resultados obtenidos se utilizarán para realizar la validación cuantitativa probabilística como: porcentaje de reconocimiento de objeto y también evaluando las métricas adecuadas para cada una de las demás características específicas del objeto, por ejemplo, para la ubicación espacial se medirá el error medio cuadrático. Cabe resaltar que con el fin de realizar una mejor

Presupuesto					
Horas nómina					
Concepto	Nombre	Escalafón	Horas mes	Sede / Seccional o Externo	Total (\$)
Horas Nomina (Investigador Principal)	Fabian Eduardo Pérez Gordillo	2	35	USTA Sede Principal	COP 9.386.563
Horas Nomina (Co-Investigadores)	Edgar Camilo Camacho Poveda	2	29	USTA Sede Principal	COP 7.777.437

FINANCIACIÓN	RECURSO	DESCRIPCIÓN	Valor partida	Valor contrapartida (Externa)	Total (COP)
RUBROS	Servicios Técnicos	N/A			\$ 0
	Salidas de campo	N/A			\$ 0
	Equipos	N/A			\$ 0
	Materiales, insumos y software	<ul style="list-style-type: none"> ▪ Materiales para el inicio del desarrollo de robot de servicio doméstico. ▪ GPU para entrenamiento de algoritmos de inteligencia artificial y aprendizaje de máquina. ▪ Materiales mecánicos y electrónicos en general. 			\$ 4.000.000
BOLSAS	Papelería	N/A			\$ 0
	Fotocopias	N/A			\$ 0
	Material bibliográfico	N/A			
	Auxilio de transporte	N/A			\$ 0
	Movilidad	Participación RoboCup 2020, en Burdeos, Francia para 2 personas.			\$ 14.000.000
	Publicaciones (Artículos, proceso editorial y traducción)	N/A			\$ 0
TOTAL DEL PROYECTO:					\$ 18.000.000

Referencias bibliográficas

- [1] L. B. a. L. I. Nizar Massouh, «RoboCup@Home-Objects: benchmarking object recognition for home robots,» *RoboCup 2019*, 2019.
- [2] SoftBank Robotics, «Pepper,» [En línea]. Available: <https://www.softbankrobotics.com/emea/en/pepper>. [Último acceso: 07 2019].
- [3] C. M. U. R. T. C. U. o. P. U. o. S. C. S. U. U. o. C.-B. U. o. W. U. o. S. C. S. U. U. o. C.-B. U. o. W. G. I. of Technology, «A Roadmap for U.S. Robotics: From Internet to Robotics,» Georgia Institute of Technology, 2013. [En línea]. Available: <https://books.google.com.co/books?id=KPhQngEACAAJ>.
- [4] S. c/o euRobotics AISBL, Robotics 2020 Multi-Annual Road-map For Robotics in Europe, 2015.
- [5] M. A. Y. K. I. N. a. E. O. H. Kitano, «Robocup: The robot world cup initiative,» *of the First International Conference on Autonomous Agents*, ser. *AGENTS*, pp. 340 - 347, 1997.
- [6] RoboCup@Home, «Team Description Papers,» [En línea]. Available: <https://github.com/RoboCupAtHome/AtHomeCommunityWiki/wiki/Team-Description-Papers>. [Último acceso: 07 2019].
- [7] E. R. F. P. S. R. S. T. M. U. M. U. M. a. C.-I. M. Carlos A. Quintero, «2019 SinfonIA Pepper Team Description Paper,» 2019.
- [8] RoboCup, «RoboCup@Home,» [En línea]. Available: <http://www.robocupathome.org/media>. [Último acceso: 07 2019].
- [9] D. H. J. R.-d.-S. K. S. T. v. d. Z. Luca Iocchi, «RoboCup@Home: Analysis and results of evolving competitions for domestic and service robots,» *Artificial Intelligence*, vol. 229, pp. 258 - 281, 2015.
- [10] T. v. d. Z. L. I. a. S. S. Thomas Wisspeintner, «RoboCup@Home: Scientific Competition and Benchmarking for Domestic Service Robots,» 2009.
- [11] Carnegie Mellon University, «Carmen, Robot Navigation Toolkit,» [En línea]. Available: (<http://carmen.sourceforge.net/>).
- [12] «The Player/Stage Project,» [En línea]. Available: (<http://playerstage.sourceforge.net/>).
- [13] «The Mobile Robot Programming Toolkit,» [En línea]. Available: <http://babel.isa.uma.es/mrpt/index.php/Main Page>.
- [14] OpenCV, «The Open Computer Vision Library,» [En línea]. Available: (<http://sourceforge.net/projects/opencv/>).
- [15] «ReadyBot,» [En línea]. Available: <http://www.readybot.com/>.

- [16] «PR2,» [En línea]. Available: <http://www.willowgarage.com>.
- [17] «Wakamaru,» [En línea]. Available: <http://www.mhi.co.jp/kobe/wakamaru/english>.
- [18] «PaPeRo,» [En línea]. Available: http://www.nec.co.jp/robot/english/robotcenter_e.htm.
- [19] Toyota, «Partner Robot FAMILY,» [En línea]. Available: https://www.toyota-global.com/innovation/partner_robot/robot/.
- [20] SoftBank Robotics, «Pepper Documentation,» [En línea]. Available: http://doc.aldebaran.com/2-4/home_pepper.html.
- [21] F. S. N. Z. R. F. a. M. C. Dong Seon Cheng, «Semantically-driven automatic creation of training sets for object recognition,» *Computer Vision and Image Understanding*, vol. 131, pp. 56-71, 2015.
- [22] A. S. I. H. G. Krizhevsky, «Imagenet classification with deep convolutional neural networks,» *Proc NIPS*, 2012.
- [23] C. L. W. J. Y. S. P. R. S. A. D. E. D. V. V. R. A. Szegedy, «Going deeper with convolutions. CoRR,» 2014.
- [24] «Yolo,» [En línea]. Available: <https://pjreddie.com/darknet/yolo/>. [Último acceso: 07 2019].
- [25] D. A. D. E. C. S. a. S. R. W. Liu, «SSD: Single shot multibox detector,» *ECCV*, 2016.
- [26] P. G. R. G. K. H. P. D. a. F. A. R. (. Tsung-Yi Lin, «Focal Loss for Dense Object Detection,» *ICCV*, nº 2980 - 2988.
- [27] A. F. Joseph Redmon, «YOLOv3: An Incremental Improvement,» 2018.
- [28] L. D.] L. Von Ahn, «Labeling images with a computer game,» *Proceedings of the SIGCHI Conference on Human factors in Computing Systems, ACM*, pp. 319 - 326, 2014.
- [29] G. W. C.-W. N. Y.-G. J. S. Zhu, «On the sampling of web images for learning visual concept classifiers,» *Proceedings of the ACM International Conference on Image and Video Retrieval, ACM*, pp. 50 - 57, 2010.
- [30] W. D. R. S. L.-J. L. K. L. L. F.-F. J. Deng, «ImageNet: a large-scale hierarchical image database,» *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 248 - 255, 2009.
- [31] M. R. M. S. J. S. S. V. J. Y. T. Dean, «Fast, accurate detection of 100,000 object classes on a single machine,» *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1814 - 1821, 2013.

- [32] ROS.org, «ROS,» [En línea]. Available: <http://wiki.ros.org/>. [Último acceso: 07 2019].
- [33] ROS.org, «vision_opencv,» [En línea]. Available: http://wiki.ros.org/vision_opencv.
- [34] F. B. T. T. J. Y. N. H. a. B. C. Nizar Massouh, «Learning Deep Visual Object Models From Noisy Web Data: How to Make it Work,» 2017.
- [35] A. K. T. M. A. T. C. Vondrick, «HOGgles: visualizing object detection features,» *IEEE International Conference on Computer Vision (ICCV)*, 2013.
- [36]]. C. F. (Ed.), «WordNet: An Electronic Lexical Database,» *MIT Press*, 1998.

