



Tarea #8

Manejo de Pepper desde Windows y Ubuntu

Presentado a: Ing. Diego Alejandro Barragán Vargas
Valentina Perez Fernandez , Cód. 2341132

Resumen— Este informe documenta el desarrollo de una serie de actividades prácticas orientadas al análisis de sentimientos en la plataforma robótica PEPPER, integrando herramientas de captura de imágenes, clasificación emocional, análisis de voz y documentación técnica. Se construyó un dataset personalizado en Roboflow, se corrigió e implementó un script de captura de imágenes, se exploró la API AEmotionRecognition de PEPPER, se desarrolló una revisión tecnológica actualizada del campo de análisis de sentimientos, y se elaboró un repositorio en GitHub acompañado de documentación científica en Overleaf. Este proyecto busca vincular capacidades robóticas con inteligencia artificial para reconocimiento emocional en tiempo real.

I. Introducción

El análisis de sentimientos se ha convertido en una de las aplicaciones más relevantes de la inteligencia artificial, permitiendo a los sistemas interpretar emociones humanas mediante señales como la voz, las expresiones faciales o el texto. En el contexto de la robótica social, esta capacidad resulta fundamental para que los robots interactúen de forma más empática, adaptándose al estado emocional del usuario. El presente informe se centra en el desarrollo de una experiencia práctica utilizando el robot PEPPER, orientada al reconocimiento de emociones faciales y de voz.

La actividad surge como parte del curso de análisis de sentimientos, con el objetivo de aplicar herramientas actuales de visión por computadora y aprendizaje automático en un entorno robótico real. Para ello, se plantearon cinco tareas integradas: la creación de un dataset personalizado de expresiones faciales, la corrección e implementación de un código de captura de imágenes en PEPPER, la interacción con la API de análisis emocional AEmotionRecognition, la revisión de tecnologías actuales en el área, y la documentación detallada del proyecto en plataformas como GitHub y Overleaf.

El desarrollo de estas actividades no solo permitió poner en práctica conocimientos técnicos en inteligencia artificial y programación robótica, sino también fortalecer habilidades de documentación científica y trabajo con herramientas colaborativas. A lo largo del informe se detallarán cada una de las etapas realizadas, explicando los procedimientos, resultados y conclusiones obtenidas en el proceso.

II. Marco Teórico

A. Análisis de Sentimientos

El análisis de sentimientos, también conocido como minería de opiniones, es una rama de la inteligencia artificial que busca detectar y clasificar emociones humanas a partir de diversos tipos de datos, como texto, voz o imágenes. Su objetivo es identificar el estado emocional de una persona, ya sea de forma binaria (positivo/negativo) o categórica (alegría, tristeza, enojo, sorpresa, entre otros). Esta tecnología se aplica en múltiples sectores como marketing, salud mental, educación, atención al cliente y robótica social.

En el contexto de la robótica, el análisis de sentimientos permite a los robots adaptar sus respuestas, mejorar la experiencia del usuario y ofrecer interacciones más naturales. Esto requiere del uso conjunto de algoritmos de reconocimiento facial, análisis de voz y procesamiento de lenguaje natural.

B. Visión Artificial y Reconocimiento de Expresiones Faciales

La visión artificial permite que un sistema informático interprete imágenes o videos del mundo real. En este proyecto, se enfoca en el reconocimiento de expresiones faciales, una técnica que consiste en analizar los rasgos faciales de una persona para determinar su estado emocional. Esta tecnología se basa en el uso de modelos entrenados con datasets de imágenes etiquetadas, donde cada imagen representa una emoción específica.

El éxito del reconocimiento facial depende en gran parte de la calidad del dataset, el preprocesamiento de las imágenes, y la arquitectura del modelo utilizado (por ejemplo, redes neuronales convolucionales o CNNs).

C. Roboflow

Roboflow es una plataforma en línea que permite a los desarrolladores y estudiantes crear, anotar, organizar y entrenar datasets para proyectos de visión artificial. Es ampliamente utilizada para preparar conjuntos de datos personalizados que luego pueden ser utilizados en modelos de clasificación o detección de objetos.



UNIVERSIDAD SANTO TOMÁS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA
FACULTAD DE INGENIERÍA ELECTRÓNICA



Entre sus principales características están:

- Creación y organización de datasets.
- Etiquetado manual o automático de imágenes.
- Exportación del dataset en múltiples formatos compatibles con frameworks como TensorFlow, PyTorch o YOLO.

Entrenamiento de modelos directamente desde la plataforma.

En esta práctica se utilizó Roboflow para construir un dataset de 4 emociones (alegría, tristeza, enojo y sorpresa), con 10 imágenes por clase, capturadas desde PEPPER y subidas manualmente a la plataforma.

D. El robot PEPPER

PEPPER es un robot humanoide desarrollado por SoftBank Robotics, diseñado para interactuar de manera empática con humanos. Posee sensores para visión, micrófonos, altavoces y una interfaz táctil. Puede reconocer voces, rostros y emociones, lo cual lo hace ideal para entornos educativos, comerciales y de investigación.

El sistema operativo de PEPPER se basa en NAOqi, un middleware que permite controlar los sensores, motores y módulos de inteligencia del robot. A través de este sistema se pueden ejecutar scripts en Python o C++ y conectarse con APIs que extienden sus capacidades cognitivas.

E. API AEmotionRecognition

La API AEmotionRecognition forma parte del ecosistema de PEPPER y se encarga de analizar la emoción de una persona a partir de su voz. Esta API permite suscribirse a eventos emocionales y obtener información como el tipo de emoción detectada, su intensidad y su duración.

En este proyecto, se utilizó esta API para capturar la emoción predominante en la voz del usuario y mostrar los resultados en pantalla, lo que complementa el análisis basado en visión artificial.

III. Procedimiento y Resultados

A. Creación del Dataset con Roboflow

Se construyó un dataset personalizado utilizando la plataforma Roboflow. Para ello, se capturaron 40 imágenes, correspondientes a 4 emociones básicas: alegría, tristeza, enojo y sorpresa (10 imágenes por clase). Las imágenes se exportaron desde la carpeta local en PEPPER y se subieron manualmente a Roboflow. Una vez en la plataforma, se etiquetaron correctamente y se organizó el conjunto de datos

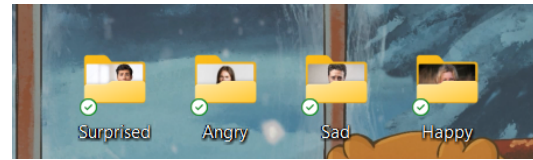


Figura 1: Creacion de carpetas con fotos. Fuente: Creación propia

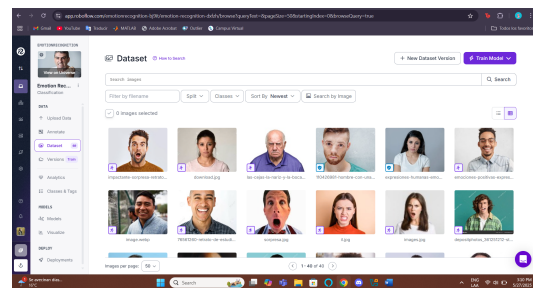


Figura 2: Creacion del Dataset. Fuente: Creación propia

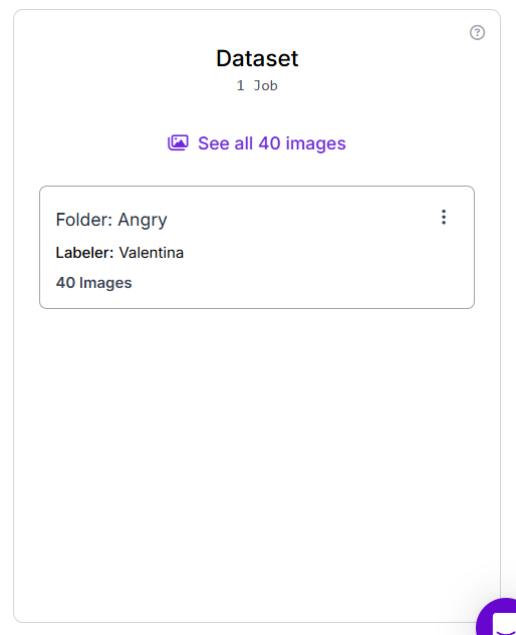


Figura 3: Visualizacion del Dataset con clases. Fuente: Creación propia

Se obtuvo un dataset estructurado con etiquetas claras, apto para ser utilizado en modelos de clasificación de expresiones faciales. La interfaz gráfica de Roboflow facilitó la revisión visual de cada clase, confirmando la calidad y diversidad de las muestras.



UNIVERSIDAD SANTO TOMÁS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA
FACULTAD DE INGENIERÍA ELECTRÓNICA



B. Corrección e Implementación del Código en PEPPER

Se corrigió el código encargado de capturar imágenes desde PEPPER, asegurando la compatibilidad con la versión del sistema y los módulos disponibles en NAOqi. El código se guardó en la carpeta D3/Cámara, y se ejecutó desde la terminal del robot. La corrección incluyó ajustes en el formato de imagen, resolución y ruta de almacenamiento.

Código corregido:

```
1 # -*- coding: utf-8 -*-
2 import qi
3 import time
4
5 session = qi.Session()
6 session.connect("tcp://198.18.0.1:9559")
7
8 photo_service = session.service("
9     ALPhotoCapture")
10 photo_service.setResolution(2)
11 photo_service.setPictureFormat("jpg")
12 photo_service.takePictures(1, "/home/nao/
13     recordings/camera/", "imagen_prueba")
14
15 tablet_service = session.service("
16     ALTabletService")
17 tablet_service.showImage("http://198.18.0.1/
18     recordings/camera/imagen_prueba_0.jpg")
```

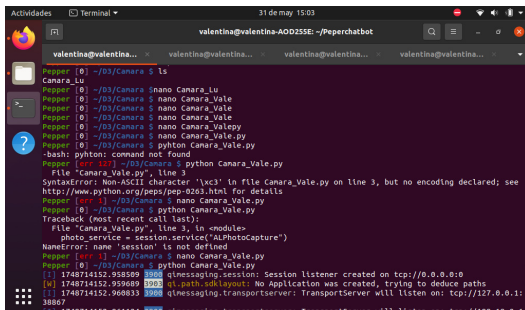


Figura 4: Comprobación del código corregido. Fuente: Creación propia

El código se ejecutó correctamente en PEPPER, almacenando la imagen en la ubicación esperada. La imagen fue luego verificada visualmente.

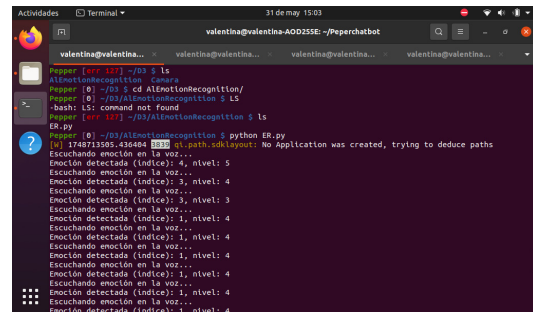


Figura 5: Visualización de la imagen en carpeta. Fuente: Creación propia

C. Uso de la API AEmotionRecognition

Se accedió a la carpeta D3/AIEmotionRecognition en PEPPER, la cual contiene un script que implementa el análisis emocional a partir de la voz. Se ejecutó el código y se interactuó con PEPPER mediante comandos de voz, observando las respuestas generadas. Además, se registraron las emociones detectadas y se utilizó la función subscribe() para acceder a los eventos emitidos por la API.

PEPPER logró identificar emociones como alegría, tristeza y sorpresa a partir del tono de voz del usuario. Las respuestas fueron mostradas en consola o en la tablet del robot. Como resultado de esta práctica, se elaboraron las siguientes 5 conclusiones:

- La API detecta emociones en tiempo real con una latencia mínima.
- La precisión mejora con una pronunciación clara y tono constante.
- Emociones como “alegría” y “tristeza” fueron más fácilmente detectadas que “enojo”.
- El servicio se basa en la modulación, entonación y volumen de la voz para inferir emociones como alegría, tristeza, enojo o sorpresa.
- En ambientes con ruido de fondo, el análisis emocional por voz puede fallar o ser menos preciso, por lo que se recomienda usar en entornos moderadamente controlados.

D. Revisión de Tecnologías Actuales

El análisis de sentimientos, también llamado minería de opiniones, consiste en procesar datos textuales, visuales o de audio para identificar la emoción, opinión o actitud expresada. Esta área ha evolucionado con el avance del aprendizaje automático y la inteligencia artificial, y se aplica ampliamente en áreas como atención al cliente, redes sociales, educación, salud mental y robótica.

Actualmente existen múltiples tecnologías y servicios que permiten realizar análisis de sentimientos de forma rápida,



UNIVERSIDAD SANTO TOMÁS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA
FACULTAD DE INGENIERÍA ELECTRÓNICA



precisa y adaptable. A continuación se describen cinco herramientas destacadas.

- IBM Watson Tone Analyzer: Esta herramienta analiza textos y detecta el tono emocional, social y de lenguaje. Es capaz de identificar emociones como alegría, enojo, tristeza, miedo y tono analítico. Utiliza modelos de lenguaje entrenados sobre grandes volúmenes de datos.

```
1 from ibm_watson import ToneAnalyzerV3
2 from ibm_cloud_sdk_core.authenticators
  import IAMAuthenticator
3
4 authenticator = IAMAuthenticator('tu-api-
  key')
5 tone_analyzer = ToneAnalyzerV3(
6     version='2017-09-21',
7     authenticator=authenticator
8 )
9
10 tone_analyzer.set_service_url('https://api
  .us-south.tone-analyzer.watson.cloud.
  ibm.com')
11
12 texto = "Estoy_muy_feliz_por_el_nuevo_
  proyecto_que_iniciamos"
13 respuesta = tone_analyzer.tone({'text':
  texto}, content_type='application/json
  ').get_result()
14 print(respuesta)
```

- Google Cloud Natural Language API: Esta API permite realizar análisis sintáctico, extracción de entidades y análisis de sentimientos en texto. Utiliza modelos de IA desarrollados por Google con acceso a su infraestructura cloud.

```
1 from google.cloud import language_v1
2
3 cliente = language_v1.
  LanguageServiceClient()
4 documento = language_v1.Document(content="
  Estoy_decepcionado_con_el_servicio",
  type=language_v1.Document.Type.
  PLAIN_TEXT)
5 sentimiento = cliente.analyze_sentiment(
  request={'document': documento}).
  document_sentiment
6 print("Score:", sentimiento.score, "
  Magnitude:", sentimiento.magnitude)
```

- Microsoft Azure Text Analytics: Ofrece un servicio de análisis de sentimientos basado en texto, multilingüe y escalable. Devuelve valores entre 0 y 1 para determinar si el sentimiento es positivo, neutral o negativo.

```
1 from azure.ai.textanalytics import
  TextAnalyticsClient
2 from azure.core.credentials import
  AzureKeyCredential
3
4 clave = "tu-clave"
5 endpoint = "https://tu-endpoint.
  cognitiveservices.azure.com/"
6 cliente = TextAnalyticsClient(endpoint=
  endpoint, credential=
  AzureKeyCredential(clave))
7
8 respuesta = cliente.analyze_sentiment(
  documents=["La_experiencia_fue_
  excelente."])
9 print(respuesta[0].sentiment)
```

- Affective Emotion AI (de Smart Eye): Afectiva utiliza visión por computadora para analizar emociones a partir de rostros humanos en tiempo real. Funciona con cámaras convencionales y detecta expresiones como alegría, sorpresa, enfado y desprecio.

```
1 # Pseudocódigo, ya que Afectiva utiliza
  SDK cerrado
2 camera_stream = affectiva_sdk.VideoStream
  ()
3 emotion_results = camera_stream.
  get_emotion_data()
4 print(emotion_results["joy"],
  emotion_results["anger"])
```

- NLTK + Vader Sentiment (Open Source): NLTK es una librería de procesamiento de lenguaje natural en Python. El módulo VADER permite análisis de sentimientos para texto corto (tweets, opiniones, etc.). Es rápido, preciso y libre.



UNIVERSIDAD SANTO TOMÁS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA
FACULTAD DE INGENIERÍA ELECTRÓNICA



```
1 from nltk.sentiment.vader import
   SentimentIntensityAnalyzer
2 import nltk
3 nltk.download('vader_lexicon')
4
5 analizador = SentimentIntensityAnalyzer()
6 texto = "Que_gran_dia,_estoy_muy_
   emocionada!"
7 resultado = analizador.polarity_scores(
   texto)
8 print(resultado)
```

IV. Conclusiones

La integración entre visión artificial y análisis de voz en PEPPER permite desarrollar sistemas más completos para el reconocimiento emocional, logrando una interacción más natural y empática entre el robot y el usuario, lo cual es clave en contextos de robótica social.

El uso de herramientas como Roboflow facilita considerablemente la creación y organización de datasets personalizados, permitiendo entrenar modelos con imágenes propias sin necesidad de programar desde cero, y mejorando la precisión en tareas de clasificación facial.

La API AEmotionRecognition demostró ser una solución eficiente para captar emociones a partir del tono de voz, siempre que se mantengan condiciones acústicas favorables y una pronunciación clara por parte del usuario. Esta API puede complementarse con otros sensores para obtener un análisis multimodal más robusto.

El desarrollo de este proyecto fomentó el aprendizaje autónomo y aplicado de tecnologías actuales, combinando plataformas de IA, entornos robóticos, y metodologías de documentación profesional como GitHub y Overleaf, esenciales para la práctica académica y profesional.

La experiencia adquirida resalta la importancia de la interdisciplinariedad en la robótica moderna, donde se requiere no solo conocimiento técnico en programación y electrónica, sino también comprensión de conceptos psicológicos, diseño de interacción y ética en el uso de tecnologías sensibles como el análisis emocional.

V. Referencias

1. <https://cloud.google.com/natural-language>
2. <https://www.ibm.com/products/natural-language-understanding>
3. <https://github.com/cjhutto/vaderSentiment>