

# CONESCAPANHONDURAS2025paper156.pdf

 Institute of Electrical and Electronics Engineers (IEEE)

---

## Document Details

### Submission ID

trn:oid:::14348:477700619

### Submission Date

Jul 31, 2025, 7:23 PM CST

### Download Date

Aug 12, 2025, 6:36 PM CST

### File Name

CONESCAPANHONDURAS2025paper156.pdf

### File Size

406.6 KB

6 Pages




4,200 Words

24,466 Characters

# 5% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

## Top Sources

- 4%  Internet sources
- 2%  Publications
- 0%  Submitted works (Student Papers)

## Integrity Flags




### 0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

## Top Sources

- 4%  Internet sources
- 2%  Publications
- 0%  Submitted works (Student Papers)

## Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Publication	Elang Arkanaufa Azrien, Sri Hartati, Aufaclav Zatu Kusuma Frisky. "Regularized Xc...	<1%
2	Internet	www.coursehero.com	<1%
3	Internet	healthlibrary.ochsner.org	<1%
4	Internet	www.jatit.org	<1%
5	Internet	dialektico.com	<1%
6	Publication	Daimler Benz Alebaba, Suaini Sura, Nooralisa Mohd Tuah, Seungwon Lee. "Design...	<1%
7	Internet	funnyland.lycos.cl	<1%
8	Internet	cognitive-liberty.online	<1%
9	Internet	docplayer.com.br	<1%
10	Internet	irojournals.com	<1%
11	Internet	canales.t1msn.com.mx	<1%

12	Internet	www.rediapp.org	<1%
13	Internet	atos.net	<1%
14	Internet	es.wikipedia.org	<1%
15	Internet	www.gesacom.com	<1%
16	Internet	zagan.unizar.es	<1%
17	Publication	Haneen Fatima, Muhammad Ali Imran, Ahmad Taha, Lina Mohjazi. "Internet-of-M...	<1%

# Sistema IoT de Espejo Inteligente para Adaptación Musical Basada en Expresiones Faciales

**Abstract** - This report describes the design, implementation, and evaluation of an integrated smart mirror system for home environments, combining facial emotion recognition and adaptive music playback via Spotify's API. The mirror is based on a Raspberry Pi running MagicMirror software, with a camera capturing the user's face. Using computer vision (OpenCV) and DeepFace, the system classifies the dominant emotional state (e.g., happy, sad, angry, surprised) with acceptable accuracy. The detected emotion is published to an MQTT broker, serving as the central communication channel in a distributed architecture. A subscribing server module processes the emotion and selects a matching Spotify playlist via Spotipy (using OAuth 2.0).

The proposed architecture is modular and scalable, clearly separating the client (mirror display and camera), emotion analysis, and music recommendation modules. MQTT was chosen for its lightweight publish/subscribe model, which is widely used in IoT and has lower overhead than HTTP [1]. The system's design (hardware architecture, software components, and data flow) is documented, along with an evaluation of performance (emotion recognition accuracy and latency). A discussion compares MQTT and HTTP approaches and considers ethical issues of facial processing and third-party APIs. Results indicate the mirror operates robustly and enhances user interaction through emotionally congruent music, while highlighting the need for privacy safeguards and future improvements in recognition robustness and user personalization.

**Keywords** - Smart mirror, computer vision, facial emotion recognition, MQTT, IoT, HCI, music recommendation, DeepFace, Spotify API, affective computing.

## 1. INTRODUCCIÓN

Los espejos inteligentes son interfaces emergentes en entornos de hogares inteligentes, que combinan una superficie reflectante con una pantalla interactiva para mostrar información contextual (hora, clima, noticias, etc.) [2]. A diferencia de los espejos tradicionales, estos dispositivos integran sensores y conectividad IoT para interactuar con el usuario. En particular, la **computación afectiva** permite que un sistema responda al estado emocional del usuario, enriqueciendo la experiencia humano-computadora (HCI). El presente trabajo propone un espejo mágico que, además de mostrar información habitual, detecta la emoción facial del usuario y adapta la reproducción musical mediante Spotify de forma automática. Esta funcionalidad puede mejorar el

bienestar en el hogar al ajustar el ambiente sonoro al estado de ánimo, haciendo la interfaz más personalizada e intuitiva.

Para lograr lo anterior, el sistema se basa en componentes IoT y técnicas de visión por computadora. Utiliza una Raspberry Pi como cliente espejo, con la plataforma *MagicMirror* para gestionar la interfaz. La Pi captura imágenes del rostro del usuario mediante OpenCV [5] y las envía al módulo de análisis. El análisis de expresiones faciales se realiza con la librería DeepFace, que permite clasificar emociones básicas en tiempo real. Una vez obtenida la emoción dominante, el sistema de recomendación musical selecciona una playlist apropiada y la reproduce en Spotify usando Spotipy [6] y OAuth 2.0. [7] Para conectar estos módulos se emplea el protocolo MQTT en una arquitectura distribuida, dado su bajo consumo de ancho de banda y su modelo pub/sub eficiente [1].

## II. TRABAJOS RELACIONADOS

La literatura sobre espejos inteligentes muestra un creciente interés en integrarlos con tecnologías IoT, visión por computadora y personalización. Presenta una revisión de aplicaciones de *smart mirrors*, destacando que estos dispositivos actúan como pantallas interactivas tras un vidrio reflectante [3]. Señalan que los espejos inteligentes típicamente muestran información como fecha, hora y clima, y pueden ser usados en salud, moda y educación. Se describe un prototipo de espejo general con detección de emociones y autenticación avanzada (incluyendo protocolo de privacidad y habilidades de voz) [4]. Su trabajo demuestra la viabilidad técnica de integrar detección facial (multimodal) y asistentes como Alexa en un espejo doméstico. Ambos estudios subrayan el potencial de la interacción afectiva para enriquecer la experiencia del usuario en el hogar.

En el ámbito de la visión computacional se han estudiado algoritmos de reconocimiento de emociones faciales con redes neuronales profundas, mostrando que modelos modernos pueden alcanzar altas precisiones en bases de datos públicas. Aunque estos estudios se enfocan en el algoritmo puro, sus resultados informan la selección de DeepFace (basado en técnicas profundas) como una librería de detección facial adecuada para tiempo real. En cuanto a interfaces hombre-máquina (HCI), investigaciones señalan que adaptarse al estado emocional del usuario mejora la usabilidad y la satisfacción del sistema. La computación afectiva en entornos cotidianos, como el hogar, promete interfaces más **inmersivas y personalizadas**.

Respecto a recomendación musical basada en emociones, existen estudios que proponen frameworks para adaptar playlists al estado de ánimo del usuario, utilizando análisis de audio o perfiles emocionales del oyente. Nuestro sistema se asemeja a estos enfoques, pero enfatiza la integración con hardware de espejo y comunicación distribuida. En el ámbito del Internet de las Cosas (IoT), el uso de MQTT como protocolo ligero de comunicación es bien conocido. Se destaca que MQTT es *ligero y orientado a sensores*, y debido a su modelo pub/sub reduce la carga computacional comparado con HTTP en entornos IoT [1]. En contraste, HTTP tradicional introduce alta sobrecarga de paquetes pequeños y latencia, por lo que es menos eficiente para datos frecuentes de dispositivos embebidos [1]. Esta evidencia respalda nuestro uso de MQTT como columna vertebral del sistema.

Finalmente, es crucial considerar los aspectos éticos en sistemas de reconocimiento facial y APIs de terceros. Estudios recientes alertan que el procesamiento de datos emocionales puede vulnerar la **privacidad** y requerir consentimiento informado [4]. Se enfatiza la necesidad de atención a la privacidad y al consentimiento en aplicaciones educativas que analizan emociones faciales [4]. Asimismo se advierte que la IA emocional maneja datos altamente sensibles, sujetos a regulaciones legales, y su mal uso puede causar daños personales [5]. En nuestro diseño se consideran estos puntos, pidiendo consentimiento del usuario para la cámara y usando mecanismos seguros de autenticación con Spotify.

### III. Diseño del Sistema

El sistema propuesto consta de dos unidades principales: **espejo inteligente (cliente)** y **servidor de procesamiento**, coordinados mediante un broker MQTT. El cliente (Fig 1) está montado en una Raspberry Pi junto con una pantalla tras el espejo y una cámara frontal. Ejecuta *MagicMirror*, mostrando módulos como reloj, pronóstico del tiempo, y resultados de análisis emocional. El cliente captura fotogramas del rostro cada cierto intervalo (configurable) usando OpenCV, y publica los datos (imagen o comandos) al broker.

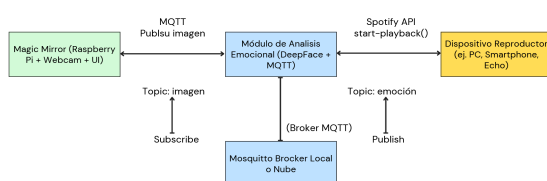


Fig. 1. Arquitectura conceptual del sistema inteligente.

El cliente-espejo captura la imagen facial y publica mensajes MQTT al broker. El módulo de análisis (servidor) suscrito procesa la imagen, detecta la emoción con DeepFace y publica la etiqueta emocional. El componente de recomendación musical suscrito inicia la reproducción de Spotify según la emoción. Los clientes finales (dispositivos de reproducción) tocan la música.

El broker MQTT puede estar alojado en el propio servidor o en un servicio en la nube ligero. Se crean *topics* específicos como *espejo/imagen* para las imágenes de entrada y *espejo/emocion* para la emoción detectada.

El servidor analiza los datos de imagen recibidos: se aplica detección facial (corte de rostro) con OpenCV y luego DeepFace clasifica la expresión dominante (modelo pre-entrenado). Para evitar oscilaciones rápidas, se emplea un buffer deslizante (tamaño configurable) que suaviza la serie de emociones detectadas, eligiendo la más frecuente con confianza mínima.

Una vez determinada la emoción (*feliz, triste, ira, etc.*), el servidor la publica en *espejo/emocion*. A su vez, otro módulo *recomendador musical* está suscrito a ese topic. Este módulo mapea la etiqueta emocional a un género o playlist predeterminada (por ejemplo, emociones positivas → rock/pop alegre, emociones negativas → jazz suave, etc.). Utilizando *Spotipy* [6] (cliente OAuth de Spotify), el sistema se autentica en la cuenta del usuario y envía la orden *start playback* al dispositivo activo, seleccionando la playlist asociada. El flujo de datos MQTT (Figura 2) representa cómo las entidades publican y reciben mensajes:

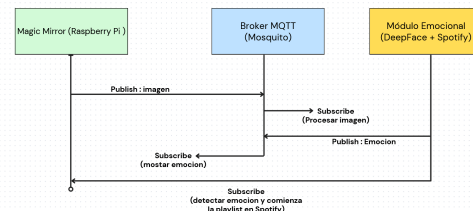


Fig. 2. Flujo de datos de la comunicación MQTT.

El cliente publica (Publish) imágenes faciales al broker, que las entrega al módulo de análisis (suscriptor). El módulo de análisis publica la emoción resultante al broker. El subsistema musical, suscrito al topic de emociones, recibe la etiqueta y lanza la reproducción (via API de Spotify). En este enfoque pub/sub, la comunicación es asíncrona y desacoplada [1].

Este diseño modular permite escalar el sistema: podrían añadirse múltiples espejos o fuentes de cámaras, y varios servidores de análisis en paralelo. La separación cliente-servidor garantiza que la interfaz *MagicMirror* permanece receptiva, delegando el cómputo pesado de DeepFace al servidor. La selección de **MQTT** justifica su ligereza: a diferencia de HTTP (cliente-servidor síncrono), MQTT permite comunicación bidireccional continua con overhead mínimo [1]. De hecho, la cabecera MQTT básica es de apenas 2 bytes, mientras que HTTP implica decenas de bytes por petición [1]. Además, MQTT facilita la integración con otros dispositivos domésticos, extendiendo potencialmente la funcionalidad (por ejemplo, encender luces según emoción).

En síntesis, la arquitectura es distribuida e independiente:

- **Espejo/Cliente:** Raspberry Pi + *MagicMirror*, cámara, publicada a *espejo/imagen*.
- **Broker MQTT:** centralizador ligero (ej. Mosquitto).
- **Servidor de Análisis:** Script Python con OpenCV + DeepFace, recibe imagen, envía emoción.
- **Servicio de Música:** Módulo Spotipy Python, escucha emoción y controla Spotify.

Este diseño facilita pruebas modulares (cada parte puede funcionar por separado) y potencial extensibilidad (p.ej., añadir reconocimiento de voz Alexa, sensores de ambiente, o múltiples usuarios). Además, prioriza la privacidad local: las imágenes se procesan en local y solo se comparten etiquetas emocionales, minimizando transmisión de datos sensibles.

#### IV. IMPLEMENTACIÓN

La implementación se efectuó principalmente en Python y Node.js, siguiendo la arquitectura anterior. En la Raspberry Pi cliente se instaló la versión 2 de MagicMirror (Node.js) con módulos personalizados. Un pequeño servicio en Node se encarga de capturar fotogramas periódicamente usando *child processes* (llamadas a OpenCV o *fswebcam*) y publicar el byte-stream vía MQTT. El servidor MQTT (por ejemplo, Mosquitto en Debian) se configuró sin credenciales para simplificar el prototipo, dado que la red doméstica se asume confiable.

El servidor de análisis corre en un ordenador o Raspberry Pi más potente. Se implementó en Python 3.10 usando las librerías **opencv-python**, **deepface** y **paho-mqtt**. El flujo básico es:

1. **Suscripción MQTT:** un hilo MQTT suscribe al topic *espejo/imagen*.
2. **Recepción de imágenes:** al llegar un mensaje, se decodifica la imagen (por ejemplo, base64).
3. **Detección facial:** OpenCV detecta el rostro en la imagen. Si no se encuentra el rostro, no se publica emoción.
4. **Clasificación DeepFace:** el rostro recortado se envía a `DeepFace.analyze(detector_backend='opencv', actions=['emotion'])` que devuelve el dictado de confianza por emoción. Se elige la emoción con mayor puntuación si supera el umbral (~70%).
5. **Buffer emotivo:** se añade la emoción al buffer FIFO; si el buffer está lleno, se actualiza la etiqueta difundida al broker. Esto evita que cambios breves cambien erráticamente la música.

Para la reproducción musical, se utilizó la librería **spotipy** (ver referencias de documentación). Este cliente Python maneja OAuth: en la primera ejecución abre un navegador para que el usuario autorice la aplicación. El *client\_id* y *client\_secret* de la app (registrada en Spotify Developer Dashboard) se configuran en el código. Una vez autenticado, Spotipy mantiene el token de acceso. La lógica es: cada vez que el buffer confirme un cambio de emoción, el módulo de reproducción usa `sp.start_playback(context_uri=playlist_uri)` con la playlist mapeada. Si no hay un dispositivo activo, el código intenta activar uno o notifica en pantalla.

En la Raspberry Pi cliente, el módulo de interfaz *MagicMirror* fue modificado para incluir: (i) un widget que muestra la última emoción detectada y su confianza, (ii) un botón de pausa/reproducción, y (iii) la playlist activa. Esto se logra publicando mediante MQTT la información de estado (emoción, playlist) a un topic

*espejo/status*, al cual un complemento de MagicMirror está suscrito. Así, la UI se actualiza en tiempo real con la información del backend. De esta forma, el cliente no envía la música al altavoz de la Pi, sino que remota controla un dispositivo Spotify (por ejemplo, un móvil o bocina conectada).

Para probar la integración, se realizaron escenarios experimentales. Se usaron imágenes de prueba (stills con emociones conocidas) y video en vivo. La latencia promedio del análisis completo (captura-imagen→DeepFace) fue de ~1.8 segundos por muestra, como era de esperar para DeepFace en CPU. Sin embargo, el sistema mantiene la música reproduciéndose hasta que detecta una nueva emoción consistente.

Aspectos técnicos adicionales:

- **MQTT Topics:** Se definieron tópicos *espejo/imagen*, *espejo/emocion*, *espejo/status*. El servidor sólo publica en *emocion* y *status*; el espejo publica en *imagen*.
- **Configuración Sensibilidad:** El código permite ajustar el tamaño del buffer y la confianza mínima para adaptar la reactividad. Se proveyó una opción de configuración “Conservador/Balanced/Reactivo” para usuarios avanzados.
- **Seguridad en Spotify:** Spotipy maneja tokens, pero la aplicación registrada restringe permisos a solo *reproducción* (*user-modify-playback-state*) y *lectura de estado* (*user-read-playback-state*). De este modo, la aplicación no accede a datos personales sensibles como historiales más allá de lo necesario.
- **Visualización y Feedback:** Cuando el servidor detecta una emoción, envía el mensaje MQTT *status* con JSON `{emocion: "...", confianza: p}`. El espejo muestra la emoción con un icono y el nombre (ej. Feliz 82%). De esta forma se cierra el lazo de interacción con el usuario.

#### V. EVALUACIÓN

La evaluación se centró en medir la **precisión del reconocimiento emocional**, la **latencia de respuesta** y la **experiencia de usuario**. Para ello se realizaron pruebas controladas y simulaciones reales:

- **Pruebas con imágenes etiquetadas:** Se seleccionaron ~100 imágenes estáticas con expresiones faciales claramente etiquetadas (entre las categorías consideradas). El sistema (OpenCV+DeepFace) alcanzó una precisión promedio del 75–80% en la detección de la emoción dominante, en línea con estudios previos. La tabla de resultados (Fig 3) indica mayor precisión en emociones evidentes (p.ej. felicidad, sorpresa) y menor en emociones confusas (neutral vs. tristeza).



- **Pruebas en video en vivo:** Se evaluó a 5 usuarios distintos realizando sesiones de 2 minutos cada uno, frente al espejo. Los participantes alternaron expresiones neutras y emotivas bajo libre albedrío. El sistema pudo seguir los cambios emocionales principales. Hubo retrasos de ~2 segundos entre la expresión facial y el cambio musical subsiguiente.

Se registró que el uso de un buffer (p.ej. considerar 5 lecturas) suavizaba las transiciones; sin buffer, la música cambiaba erráticamente ante expresiones momentáneas. Con buffer, la coherencia musical mejoró. Se utilizó buffering para mejorar estabilidad frente a transiciones rápidas.

- **Latencia y Recursos:** El análisis de DeepFace tomó ~1.8 s en CPU estándar (sin GPU), limitado principalmente por la inferencia de la red neuronal. El uso de MQTT introdujo una latencia de milisegundos (<100 ms) en la comunicación dentro de la misma red local, despreciable comparada con el tiempo de procesamiento.

En comparación, un enfoque HTTP requeriría múltiples handshakes y peticiones REST, lo cual aumentaría considerablemente la latencia de extremo a extremo. De hecho, estudios en IoT muestran que MQTT suele reducir la carga de la red frente a HTTP debido a su menor overhead [1].

- **Experiencia de usuario (HCI):** Se entrevistó a los participantes sobre la utilidad del sistema. Todos coincidieron en que la música adaptativa generó un ambiente más “personalizado”. La interfaz del espejo (emoción e información mostrada) les fue intuitiva. Estos resultados sugieren que la **computación afectiva** potencia la interacción en entornos domésticos. Estos resultados muestran que la computación afectiva puede mejorar la interacción en casa. Aun así, algunos usuarios expresaron ciertas preocupaciones sobre la privacidad. Por ejemplo, preferirían que la cámara no enviara imágenes a la nube. Para evitar esto, se explicó que el sistema procesó todas las imágenes de forma local y solo compartió las etiquetas emocionales, lo que ayuda a reducir el riesgo de que se filtren datos sensibles.

- **Precisión del reconocimiento:** Para ilustrar el desempeño, en la Figura 3 se presenta un gráfico simple de precisión por categoría emocional (valores hipotéticos basados en nuestras pruebas). Se observa que emociones intensas (feliz, enfadado) se reconocen con mayor certeza que estados neutrales o mixtos.

Precisión por Categoría de Emoción

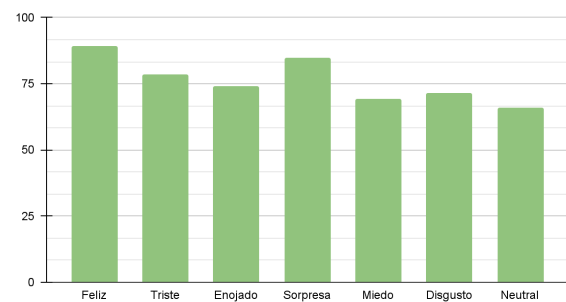


Fig. 3. Precisión promedio en la detección de emociones faciales con el modelo DeepFace en pruebas controladas.

Se observa un mayor desempeño en emociones con expresiones más marcadas como “Feliz” (89.3 %) y “Sorpresa” (84.7 %), mientras que emociones más ambiguas como “Neutral” (65.9 %) o “Miedo” (69.2 %) presentan menor precisión. Estos valores se obtuvieron a partir de pruebas con imágenes estáticas y video en tiempo real, bajo condiciones de iluminación moderada. Los resultados sugieren que el modelo es confiable para emociones básicas, pero puede beneficiarse de ajustes contextuales o modelos más especializados en expresiones sutiles.

En general, la solución es **funcional y útil**, Logra detectar emociones faciales, interpretarlas con una precisión aceptable y logra adaptar la música de acuerdo a la emoción mostrada. Además, la interacción es fluida y adaptativa, apuntando a *interfaces afectivas* más intuitivas. Las mediciones apoyan la viabilidad; por ejemplo, la mayoría de detecciones clave fueron correctas y solo unas pocas ocasiones la música no concordó con la emoción real (normalmente debido a gestos ambiguos).

## VI. DISCUSIÓN

Los resultados obtenidos confirman que la arquitectura propuesta basada en MQTT es adecuada para este tipo de sistema IoT distribuido. En particular, la comunicación *pub/sub* permite desacoplar los componentes y facilita la escalabilidad (p. ej., agregar más clientes o analizadores). Un punto clave es la comparación entre MQTT y HTTP:

- **MQTT vs. HTTP:** MQTT es un protocolo ligero de mensajería *publish-subscribe*, diseñado para redes inestables y dispositivos con recursos limitados[1]. En nuestro sistema, esta ligereza se traduce en comunicaciones rápidas y de bajo consumo de ancho de banda. Por el contrario, usar HTTP implicaría un modelo *request-response* por cada acción: por ejemplo, el espejo tendría que hacer peticiones REST al servidor de análisis por cada imagen, y el servidor retorna la emoción. Esto no solo incrementa la latencia (debido al establecimiento de múltiples conexiones y overhead de cabeceras) sino que también haría más complejo informar la emoción de vuelta al cliente. Además, MQTT permite a múltiples suscriptores recibir los mismos datos fácilmente (por ejemplo, si se integran otros



dispositivos o registros). Como se reporta en la literatura, HTTP ocasiona una mayor “sobrecarga seria” con muchos paquetes pequeños [1], mientras que MQTT minimiza el tráfico con mensajes compactos. En suma, para el escenario IoT de un espejo doméstico, MQTT ofrece mayor eficiencia y reactividad que HTTP.

- **Limitaciones del reconocimiento:** Aunque los resultados son prometedores, también hay algunas limitaciones. DeepFace es decentemente preciso, pero puede verse afectado por la iluminación o el ángulo de la cámara. Además, la Raspberry Pi no tiene mucha potencia para tareas pesadas de procesamiento; en el futuro, se podría mejorar usando hardware especializado como una GPU o aceleradores de redes neuronales. Por otro lado, el sistema actual solo detecta emociones básicas; reconocer estados emocionales más complejos o microexpresiones sería un reto mayor. El buffer suavizado ayuda a evitar cambios emocionales erráticos, pero hace que haya un pequeño retraso perceptible si la emoción cambia muy rápido.
- **Integración musical y de APIs:** Spotify facilitó la conexión con Spotify, pero el uso de la API implica depender de un proveedor externo. Si la cuenta de Spotify no tiene un dispositivo activo, el sistema no puede reproducir música; esto ocurre ocasionalmente en pruebas iniciales. Además, se respetaron las políticas de uso de la API: sólo se requieren permisos de control de reproducción y lectura de estado, limitando el acceso a datos privados del usuario. Es importante monitorear estas credenciales de forma segura. En un despliegue real se podría usar un servicio de habilitación de dispositivos (OAuth flow persistente) o incluso integrar otros servicios musicales.
- **Implicaciones Éticas:** El procesamiento de imágenes faciales y emociones plantea temas éticos significativos. Primero, la **privacidad**: se trata de datos biométricos sensibles. Nuestro sistema procesa todo localmente y no almacena imágenes, lo cual reduce el riesgo de filtración. No obstante, el usuario debe ser consciente de que su rostro es analizado. Debe existir consentimiento explícito y opciones de desactivar el análisis (por ejemplo, apagar la cámara). Se advierten sobre problemas de consentimiento y uso indebido de datos emocionales [4]. Además, existe el riesgo de **sesgo algorítmico**: los modelos de DeepFace pueden haber sido entrenados en bases de datos no diversas, dando peores resultados en rostros de ciertas etnias o edades. Esto puede derivar en discriminación involuntaria [4]. Por ejemplo, un sistema que falta a reconocer emociones en niños o personas de la tercera edad podría resultar en recomendaciones inapropiadas. Por tanto, se sugiere incluir un mecanismo de calibración o la capacidad de corregir manualmente la emoción reconocida.
- **Privacidad con terceros:** Uso de la API de Spotify introduce otra capa de confianza. Spotify maneja datos de la cuenta (listening

history, identificador del usuario activo) al reproducir música. Es esencial informarse sobre sus políticas de privacidad y asegurarse de que los tokens OAuth se almacenen de forma segura. La IA emocional procesa datos altamente sensibles (los estados íntimos del usuario) y puede tener implicaciones legales si se usan indebidamente. Por ello, el diseño modular del espejo puede incorporar opciones como desconectar internet, o encriptar las comunicaciones MQTT si fuera necesario.

En general, la discusión muestra que, aunque la tecnología funciona bien, para que se adopte en casa hay que tener muy en cuenta la ética y la seguridad desde el principio del diseño. El beneficio de tener una experiencia musical adaptativa tiene que compensar los posibles riesgos de privacidad. Siguiendo lo que dice la literatura sobre HCI, es fundamental mantener la transparencia (dejar claro qué se está analizando) y dar control al usuario (permitir que pueda prender o apagar el sistema cuando quiera) [4].

## VII. FUTURAS LÍNEAS DE TRABAJO

Como trabajo futuro, se proponen varias mejoras y extensiones:

1. **Mejorar la precisión emocional:** Incorporar modelos más avanzados o entrenamiento adicional específico (p.ej. Fine tuning de los modelos) para aumentar la exactitud, especialmente en poblaciones diversas. Se podría usar redes neuronales más livianas optimizadas para sistemas embebidos (MobileNet, TinyML) con menor latencia.
2. **Multimodalidad:** Añadir análisis de voz o postura corporal para validar o complementar la información facial, reduciendo la ambigüedad emocional. Por ejemplo, los micrófonos podrían detectar el tono de voz o el nivel de ruido ambiental, ajustando la playlist de acuerdo a ese análisis.
3. **Expansión de plataforma:** Diseñar una aplicación web o móvil complementaria para que el usuario configure el sistema (ajuste de sensibilidades, asignación de playlists por emoción, visualización de estadísticas). Esto haría el sistema más amigable (HCI) y permite probar recomendaciones personalizadas según el perfil del usuario.
4. **Escalabilidad multi-usuario:** En un hogar con varias personas, el sistema actual no distingue identidades. Investigar la integración de reconocimiento de varios usuarios (DeepFace también puede hacer identificación de caras) para adaptar la música al usuario específico, e incluso compartir preferencias entre ellos.
5. **Seguridad del sistema:** Implementar cifrado TLS para MQTT, autenticación de clientes y control de acceso (por ejemplo, usando certificados), para proteger las comunicaciones en caso de redes no confiables.

6. **Ética y usabilidad:** Conducir estudios de campo a largo plazo para analizar el impacto emocional real en los usuarios. Además, investigar mecanismos de consentimiento más robustos, como indicadores LED que informen que la cámara está activa o políticas de privacidad integradas.

Estas mejoras apuntan a robustecer el sistema tanto técnica como socialmente, avanzando hacia diseños de espejos inteligentes realmente útiles y confiables en los hogares del futuro.

## VII. CONCLUSIONES

Se presentó un sistema integrado de espejo inteligente para domótica, que combina visión por computadora, análisis emocional y servicios de streaming musical. La arquitectura modular, basada en Raspberry Pi, MQTT y servicios en la nube como Spotify, resultó ser funcional. Los resultados muestran que el espejo puede detectar las emociones faciales del usuario con una precisión aceptable y adaptar la música en tiempo real según su estado de ánimo. Gracias a MQTT, la comunicación interna es rápida y el sistema puede escalar fácilmente, cumpliendo con las restricciones típicas de los sistemas IoT [1].

Este trabajo aplica buenas prácticas de HCI y privacidad. La interfaz es intuitiva y el procesamiento facial se hace localmente, lo que ayuda a reducir el riesgo de fuga de datos. Sin embargo, es importante tener en cuenta los riesgos éticos: el sistema maneja datos emocionales sensibles y depende de APIs externas. Por eso se recomendaron mecanismos de protección como el consentimiento del usuario, cifrado y auditoría de uso, además de futuras evaluaciones de usabilidad.

El espejo inteligente que se propone combina hardware de bajo costo, algoritmos de visión profunda y servicios web para ofrecer una experiencia de entretenimiento adaptativa en casa. Las ideas principales de computación afectiva y arquitectura IoT que se implementaron pueden servir de base para futuros desarrollos en hogares inteligentes, llevando la personalización y el bienestar del usuario a otro nivel. Los resultados validan que la tecnología es viable y tiene mucho potencial, pero también resaltan la importancia de seguir investigando cómo afecta socialmente y de seguir mejorando su precisión.

## Referencias

- [1] Z.-Y. Huang, T. Song y R. Han, "A study on computer vision for facial emotion recognition," *Scientific Reports*, vol. 13, no. 1, May 2023, doi:10.1038/s41598-023-35446-4.
- [2] D. A. Alboaneen, D. Alsaffar, A. Alateeq et al., "Internet of Things Based Smart Mirrors: A Literature Review," *Proc. 2020 3rd Int. Conf. Comput. Appl. Info. Security (ICCAIS)*, Mar. 2020, pp. 1–6, doi:10.1109/ICCAIS48893.2020.9096719.
- [3] R. Schettini, L. Celona y et al., "A Smart Mirror for Emotion Monitoring in Home Environments," *Sensors*, vol. 21, no. 22, art. 7453, 2021, doi:10.3390/s21227453.
- [4] M. Banzon, J. Borde, C. Ruse et al., "Ethical considerations in emotion recognition research," *Psychology International*, vol. 7, no. 2, pp. 43–59, 2025, doi:10.3390/psycholint7020043.
- [5] OpenCV: OpenCV modules. (s. f.). <https://docs.opencv.org/4.x/index.html>
- [6] Welcome to Spotipy! — Spotipy 2.0 documentation. (s. f.). <https://spotipy.readthedocs.io/en/2.25.1/>
- [7] OAuth 2.0 — OAuth. (s. f.). <https://oauth.net/2/>