

La Realidad Aumentada (Augmented Reality, AR) y su capacidad para integrar contenido sintético sobre una imagen real proporciona un valor incalculable en diversos campos; no obstante, la industria es uno de estos campos que más se puede aprovechar de ello. Como tecnología clave en la evolución hacia la Industria 4.0 y 5.0, la AR no solo complementa sino que también potencia la interacción humana con los procesos industriales. En este contexto, la AR se convierte en una herramienta esencial que no sustituye al factor humano, sino que lo enriquece, ampliando sus capacidades y facilitando una colaboración más efectiva entre humanos y tecnología. Esta integración de la AR en entornos industriales no solo mejora la eficiencia y precisión de las tareas, sino que también abre nuevas posibilidades para la expansión del potencial humano.

Existen numerosas formas en las que el ser humano interactúa con la tecnología, siendo la AR uno de los paradigmas más innovadores respecto a cómo los usuarios acceden a la información; sin embargo, es crucial reconocer que la AR, por sí misma, tiene limitaciones en cuanto a la interpretación del contenido que visualiza. Aunque en la actualidad podemos acceder a diferentes librerías que utilizan algoritmos para realizar una detección de imágenes, objetos, o incluso entornos, surge una pregunta fundamental: ¿hasta qué punto puede la AR comprender el contexto de lo que ve? Esta cuestión se vuelve especialmente relevante en entornos industriales. ¿Puede la AR discernir si una máquina está funcionando correctamente, o su rol se limita a la presentación de indicadores digitales superpuestos? La respuesta a estas cuestiones subrayan tanto el potencial como los límites de la AR, impulsando la búsqueda de innovaciones que permitan una mayor comprensión contextual y adaptabilidad a situaciones específicas dentro de la industria.

En el núcleo de esta tesis yace el objetivo de no solo dotar a la AR de una "inteligencia semántica" capaz de interpretar y adaptarse al contexto, sino también de ampliar y enriquecer las formas en que los usuarios interactúan con esta tecnología. Este enfoque se orienta particularmente a mejorar la accesibilidad y la eficiencia de las aplicaciones de AR en entornos industriales, que son por naturaleza restringidos y complejos. La intención es ir un paso más allá de los límites tradicionales de la AR, proporcionando herramientas más intuitivas y adaptativas para los operadores en dichos entornos.

La investigación se despliega a través de tres artículos de investigación, donde se ha desarrollado y evaluado una arquitectura multimodal progresiva. Esta arquitectura integra diversas modalidades de interacción usuario-tecnología, como el control por voz, la manipulación directa y el feedback visual en AR. Además, se incorporan tecnologías avanzadas basadas en modelos de aprendizaje automática (Machine Learning, ML) y aprendizaje profundo (Deep Learning, DL) para extraer y procesar información semántica del entorno. Cada artículo construye sobre el anterior, demostrando una evolución en la capacidad de la AR para interactuar de manera más inteligente y contextual con su entorno, y resaltando la aplicación práctica y los beneficios de estas innovaciones en la industria.

En el primer artículo, se presenta y evalúa una arquitectura compuesta por cuatro capas fundamentales: la capa de interacción, la capa de negocios, la capa física de AR y la capa semántica. Esta arquitectura se ve ampliada en los artículos subsiguientes. La evaluación de esta arquitectura demuestra su capacidad para adquirir y analizar información visual del entorno, centrándose en elementos como un botón de encendido/apagado o una válvula de presión. Para llevar a cabo la evaluación del sistema, se emplean modelos clásicos de regresión y clasificación, así como redes neuronales convolucionales (Convolutional Neural Networks, CNN). Se desarrolla una aplicación de guía paso a paso para operarios de planta, dividiéndolos en dos grupos: *AR standard application* y *AR application with semantic layer*. Además, este sistema se beneficia de la inclusión de

Transformers, una arquitectura altamente especializada en el procesamiento de información textual. Esto posibilita que el usuario realice preguntas en lenguaje natural mediante tecnologías de reconocimiento de voz, obteniendo respuestas generadas en función de la información documental disponible, como por ejemplo, documentos técnicos sobre una máquina específica. Una característica esencial que diferencia esta propuesta de otros casos de uso de AR en la industria es la posibilidad de verificar, mediante la capa semántica, las acciones que acomete el usuario (como podrían ser activar un botón o verificar el nivel de presión de una máquina) antes de continuar con el siguiente paso. De este modo, es posible reducir la carga cognitiva del operario respecto a aplicaciones de AR tradicionales.

El segundo artículo avanza en la línea de investigación del primero profundizando en el uso de la capa semántica junto a la AR, añadiendo la posibilidad de realizar una guía al usuario en un entorno más complejo mediante técnicas de localización y mapeo simultáneo (Simultaneous Localization And Mapping, SLAM). Este estudio se distingue por su enfoque de evaluación más amplio, involucrando a tres grupos distintos: uno con AR y capa semántica con interacción en lenguaje natural, otro con una aplicación de AR "ciega", y un tercero sin asistencia tecnológica, dependiendo únicamente de la documentación técnica. Esta estructura de evaluación múltiple permite una comparación exhaustiva y revela claramente las ventajas de la arquitectura propuesta. Los resultados subrayan un incremento notable en la comodidad y seguridad, demostrando cómo la capa semántica no solo mejora la interacción usuario-máquina, sino que también valida y optimiza la ejecución de tareas en entornos industriales.

Finalmente, en el tercer artículo se analiza cómo el conocimiento del experto en la materia (Subject Matter Expert, SME) puede ser aprovechado gracias a las tecnologías anteriormente mencionadas y los modelos de lenguajes masivos (Large Language Models, LLMs). La investigación destaca por su enfoque en combinar, no solo la documentación técnica existente sobre la maquinaria y los diferentes procesos, si no junto a esto, el conocimiento y experiencia del SME en forma de "píldoras" ancladas a posiciones concretas en el entorno. De este modo, el operario bien puede realizar consultas en lenguaje natural de cualquier elemento, bien puede consultar las anotaciones del SME en un punto en concreto. En todo momento, la información de la documentación técnica y el conocimiento experto son usados para devolver una respuesta al operario. La evaluación del sistema se llevó a cabo con dos grupos de usuarios que deben realizar una serie de tareas. Mientras que el grupo A tenía único acceso a documentación técnica y a un SME, el grupo B disponía de acceso a la aplicación desarrollada. Teniendo en cuenta las limitaciones que supone el disponer de un SME en todo momento en planta, la evaluación del sistema reveló una preferencia marcada por la capacidad de acceder a información experta anclada, resaltando la ventaja de disponer de esta asistencia experta inmediata, así como una solución práctica y eficiente para favorecer la transferencia de conocimiento experto en la industria.

A lo largo y tras el desarrollo de esta tesis, se han extraído las siguientes conclusiones:

- La integración de capas semánticas y LLMs con la AR mejoró significativamente la eficiencia de las tareas, especialmente en tareas complejas y cognitivas, permitiendo resultados más rápidos y precisos.
- Los sistemas de AR con capa semántica no solo mejoraron tareas complejas sino también tareas más sencillas a través de validación automática y guiado, aumentando la eficiencia general.
- La AR mejorada con IA y características de procesamiento de lenguaje natural (Natural Language Processing, NLP) condujo a un acceso más rápido a la

información y a una toma de decisiones más eficiente que los métodos tradicionales como la consulta de manuales o expertos.

- Los usuarios de sistemas de AR con capas semánticas experimentaron un aumento en la satisfacción y facilidad de uso.
- El uso efectivo de sistemas de AR en entornos complejos variados, como laboratorios textiles y plantas industriales, demostró su amplia aplicabilidad en la industria.
- Las aplicaciones de AR eran fáciles de usar, permitiendo incluso a inexpertos operar maquinaria desconocida de manera efectiva.