

**Vol. 8, Num.2**

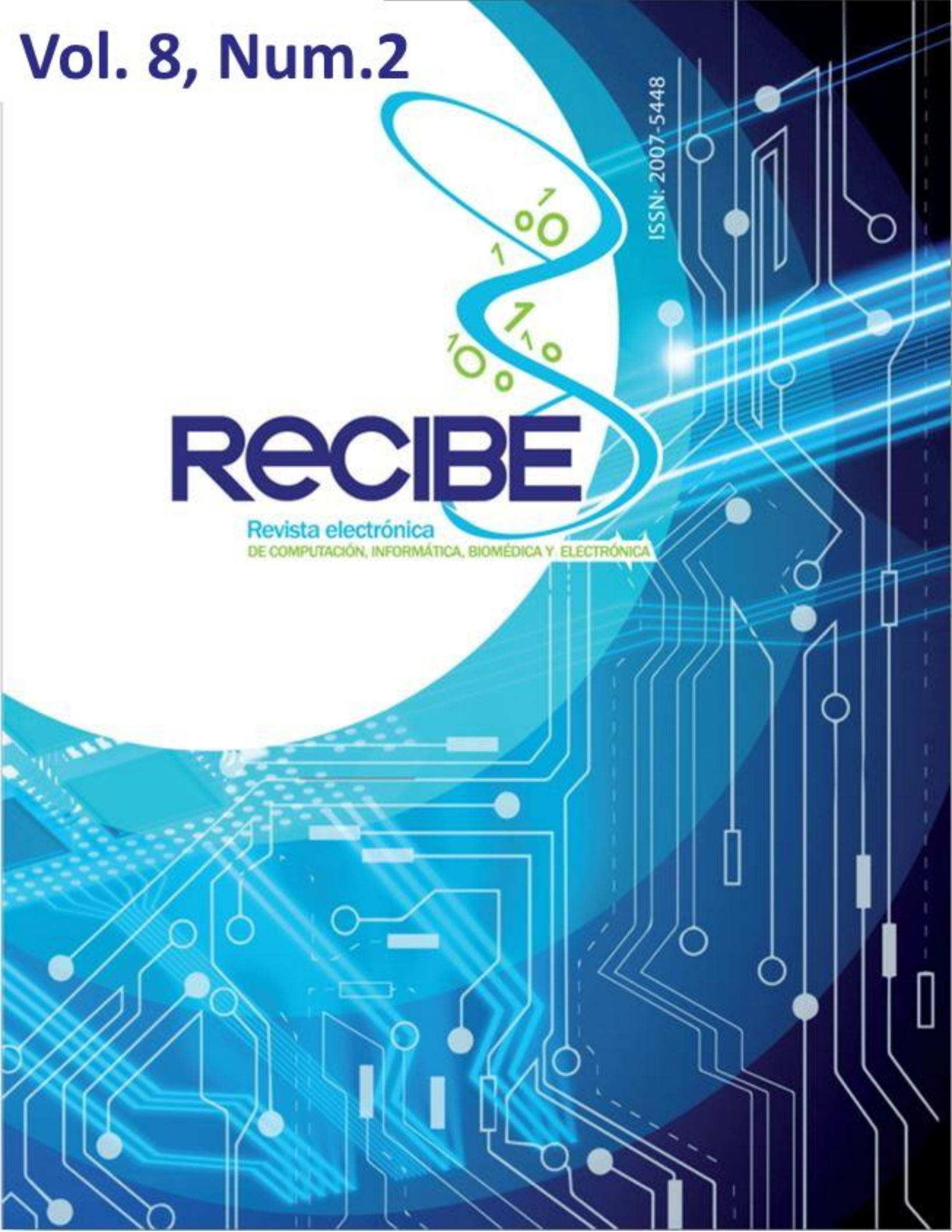
ISSN: 2007-5448

100  
110  
100

# RECIBE

Revista electrónica

DE COMPUTACIÓN, INFORMÁTICA, BIOMÉDICA Y ELECTRÓNICA



## Índice

### Computación e Informática

- Método de calificación de atributos de calidad en productos software médico regulado usando proceso analítico jerárquico y rubrica de evaluación  
Jonathan Venegas Sandoval, Francisco Javier Álvarez Rodríguez C1
- Arquitectura de un juego serio inteligente basado en retos de matemáticas básicas  
Carlos Enrique Flores Méndez, Maricela Quintana López, Héctor Rafael Orozco Aguirre, Ivonne Rodríguez Pérez C2
- SDR y GNU Radio como plataforma para un laboratorio de comunicaciones digitales  
Rogelio Jaimes Rico, Saul Lazcano Salas C3
- Análisis de sentimientos en videojuegos  
Daniel Iván Angeles Gómez, Maricela Quintana López C4
- Aplicación de Algoritmos de Votación en la Clasificación de Textos  
Saturnino Job Morales Escobar, María Esther Guevara Cruz, Héctor Rafael Orozco Aguirre C5
- Análisis del desempeño de redes neuronales profundas para segmentación semántica en hardware limitado  
Oscar Alejandro Soto Orozco, Alma Delia Corral Saenz, Claudia Elizabeth Rojo González, Juan Alberto Ramirez Quintana C6

# **Biomédica**

Development of an EEG signal analysis Application through  
a convolution of a complex Morlet wavelet: preliminary results

B1

José Humberto Trueba Perdomo,  
Ignacio Herrera Aguilar, Francesca Gasparini

# **MÉTODO DE CALIFICACIÓN DE ATRIBUTOS DE CALIDAD EN PRODUCTOS DE SOFTWARE MÉDICO REGULADO USANDO PROCESO ANALÍTICO JERÁRQUICO Y RÚBRICA DE EVALUACIÓN**

## **QUALIFICATION METHOD OF QUALITY ATTRIBUTES IN MEDICAL SOFTWARE PRODUCTS REGULATED BY THE HIERARCHICAL ANALYTICAL PROCESS AND THE EVALUATION RUBRIC**

Jonathan Venegas-Sandoval<sup>1</sup>  
*Jonathan.venegas.sandoval@gmail.com*

Francisco J. Álvarez-Rodríguez<sup>2</sup>  
*fjalvar.uaa@gmail.com*

<sup>1</sup>*Universidad Juárez Autónoma de Tabasco  
Avenida Universidad S / N Villahermosa, Tabasco, México*

<sup>2</sup>*Avenida Universidad # 940  
Ciudad Universitaria, Aguascalientes, Ags, México*

## **RESUMEN**

Las aplicaciones móviles que diagnostiquen enfermedades o funcionen como equipo médico pueden poner en peligro la salud de las personas cuando no funcionan adecuadamente. El control sobre la calidad de las aplicaciones en Google Play se basa en la usabilidad a partir de calificaciones emitidas por los usuarios. En App Store, los controles para las aplicaciones son más amplios, sin embargo, no validan por parte de expertos las aplicaciones que realizan funciones de dispositivos médico. En el presente documento se muestra una propuesta de calificación de aplicaciones móviles para el área médica utilizando una lista de atributos de calidad extraídos de normativa regulatoria. El método presentado consiste en complementar el Proceso Analítico Jerárquico (AHP por sus siglas en inglés) con una rúbrica de evaluación para determinar un grupo de aplicaciones móviles médicas que cumpla con una lista de atributos de calidad. Como caso experimental se evaluaron cinco aplicaciones móviles médicas disponibles en Google Play que miden el ritmo cardíaco con alta tasa de descargas, la menos descargada con 10,600 y una calificación de 4,5 estrellas (en un rango de 0 a 5 estrellas máximas posible), la más descargada con 198,633 descargas y una calificación de 4 estrellas.

## **PALABRAS CLAVE**

Calidad de Software, Proceso Analítico Jerárquico, Calificación de Software

## **ABSTRACT**

Mobile applications that diagnose diseases or devices such as medical equipment can endanger people's health when they have no problems. The control over the quality of applications in Google Play is based on usability according to the ratings issued by users. In the App Store, the controls for the applications are broader, however, they do not validate by experts the applications that operate functions of medical devices. This document shows a proposal for the qualification of mobile applications for the medical area using a list of quality attributes extracted from regulatory regulations. The method presented consists of complementing the hierarchical analytical process (AHP) with an evaluation publication to determine a group of mobile medical applications that includes a list of quality attributes. As an experimental case, five mobile applications available on Google Play were evaluated, which measured heart rate with a high download speed, the lowest downloaded with 10,600 and a 4,5 star rating (in a range of 0 to 5 possible maximum stars), the most downloaded with 198,633 downloads and a 4 star rating.

## **KEYWORDS**

Software Quality, Hierarchical Analytical Process, Software Qualification

## 1. INTRODUCCIÓN

Los productos de software móvil médico deben ser analizados para determinar confiabilidad y calidad debido a los riesgos potenciales para los usuarios. Instituciones e investigadores verifican y validan de forma aleatoria este tipo de aplicaciones móviles encontrando problemas de confiabilidad, seguridad y control en estos productos:

“En algo más de 1 de cada 5, sus responsables afirmaban que podían curar trastornos médicos... Cerca del 43% usaban dispositivos del propio terminal o sonidos como medios para tratamientos, sin ser dispositivos médicos” (García y Martínez, 2016, p. 30)

El método usado principalmente para determinar la calidad de las aplicaciones de salud se enfoca en analizar los productos ofertados y analizar los niveles de calidad y el posible peligro que puede representar para la salud (Fernandez y Ramos, 2014, p.100). Es el caso de MYHEALTHAPPS desarrollada por Patient View un portal británico que valida y verifica aplicaciones móviles médicas de todo el mundo basándose en las experiencias de uso y clasificándolas por tipo e idioma entre otras, sin embargo:

“La mayoría de las Apps de salud carecen de información sobre su autoría.... Los resultados de diversos estudios han revelado que menos del 35% de las Apps de salud tienen la participación de expertos médicos durante su desarrollo y que el 86% de las Apps revisadas sobre control del dolor no tenían participación sanitaria profesional en su desarrollo” (García y Martínez, 2016, p. 30)

Cuando existe riesgo para la salud de la población las instituciones gubernamentales regulan y controlan productos o mercados, en Colombia la división de gobierno encargada de esta labor es el Instituto Nacional de Vigilancia de Medicamentos y Alimentos (INVIMA) la cual como su nombre indica regula los medicamentos, alimentos y productos que puedan poner en riesgo la salud de las personas, para el caso de tecnología regula los dispositivos médicos pero no el software médico o las aplicaciones móviles que convierten un teléfono inteligente en un dispositivo médico (ISO 13485, 2016 p.18). En México se encuentra la Comisión Federal para la Protección contra Riesgos Sanitarios (COFEPRIS) la cual otorga licencias para el uso, transporte, almacenamiento y venta de productos regulados. Cuando una persona desea vender un producto en una farmacia debe tener una licencia de la COFEPRIS para verificar que el almacenamiento, manipulación y condiciones en las que vende los productos regulados no atentan contra la salud. Teniendo en cuenta que la oferta de aplicaciones para el ámbito médico aumenta constantemente (García y Martínez, 2016, pp. 29-30) y que el funcionamiento incorrecto de una aplicación médica puede generar una afectación de salud, alguna división gubernamental mexicana debería controlar el mercado de las Apps de salud. En otros países algunas instituciones como la agencia de administración de drogas y alimentos de los Estados Unidos FDA (por sus siglas

en inglés) se han adaptado a estos nuevos mercados para regular las aplicaciones móviles que puedan poner en peligro la vida de seres humanos. Esta agencia tiene amplia reputación y experiencia en la protección de la salud humana regulando drogas, alimentos y para este caso en particular, aplicaciones de software médico, la experiencia que posee regulando medicamentos, alimentos y dispositivos médicos la trasladó a las aplicaciones móviles, lo que garantiza que el control es exhaustivo y eficiente debido a su reputación, pero la agencia no enfoca sus esfuerzos en todas las aplicaciones móviles médicas. En España la Agencia de Calidad Sanitaria de Andalucía presenta la Estrategia de calidad y seguridad en aplicaciones móviles de salud, con el objetivo de brindar una serie de recomendaciones para desarrollar aplicaciones médicas de cualquier tipo, garantizando calidad y pertinencia en los productos. Esta agencia brinda un distintivo que identifica a la aplicación como una aplicación de calidad. Por ello se busca que las aplicaciones móviles médicas sean seguras y confiables, que estén en condiciones de realizar la actividad por la cual fueron desarrolladas (García y Martínez, 2016, p. 32). Sin embargo, no existe una herramienta para calificar Apps de la salud, la elección de software móvil por parte de un usuario potencial se basa en la calificación de otros usuarios a partir de la usabilidad. Sin embargo, dicha calificación no es técnica, se basa en la percepción de éxito o afinidad con el diseño y estructura de la aplicación móvil, lo que da una idea errada de calidad en la aplicación. Si a lo anterior se agrega que la usabilidad tiene diferente interpretación en otras ramas de la ciencia (Muñoz, Piattini, y Moraga, 2010 p. 62), resulta poco claro el termino para determinar la calidad de una app.

Para poder confiar la salud de un usuario a una aplicación de salud es necesaria una calificación más allá de la usabilidad, se necesita una calificación técnica avalada por expertos del área que siga una lista de atributos definida (Redrován, Loja, Correa y Piña 2017 pp. 94-113). Este método beneficia a pacientes, familiares y personal médico; sin embargo, falta establecer bajo qué condiciones se realizará y quién realizará la calificación de las aplicaciones. Esta propuesta presenta una herramienta que podría ser un complemento a las normas de calidad de productos de software o una herramienta propuesta para calificar aplicaciones en las diferentes plataformas de distribución digital. La propuesta para calificar aplicaciones móviles médicas según el nivel de calidad se basa en una lista de atributos que toda aplicación móvil debe tener según las normas ISO de calidad de software (Muñoz, Piattini, y Moraga, 2010 p. 92), donde un atributo es una característica de calidad que permite a los usuarios aceptar un producto (Fernández y Ramos, 2014 p. 99) y a su vez determina cuando una aplicación móvil puede ser utilizada sin poner en riesgo la salud del usuario.

Este trabajo se dividió en tres secciones, la primera consistió en determinar elementos considerados como críticos para la calidad de un producto de software móvil en el área médica, tomando como fuente a las instituciones certificadoras y reguladoras, incluyendo, teoría básica de validación y verificación de sistemas computarizados. Sin embargo, a pesar de que una gestión efectiva de tecnología y del proceso de desarrollo impacta en la calidad de los productos de software (Tuya,

Ramos, y Dolado, 2007 p.69), las instituciones reguladoras fijan sus objetivos en la documentación y en el producto, más que en el modelo de proceso o de gestión utilizado, por ello, estos aspectos están fuera del alcance del presente estudio. El segundo paso consistió en utilizar la herramienta AHP diseñada por Thomas L. Saaty en la década de los noventa para calificar aplicaciones móviles. Complementando dicho método se propone una rúbrica de evaluación como métrica para calificar las aplicaciones y determinar cuándo pueden ser recomendadas al usuario utilizando el rango de evaluación del AHP más el número cero. Finalmente, se muestra una prueba de aplicación realizada a cinco aplicaciones móviles concernientes al monitoreo cardíaco.

## 2. MÉTODO

Las herramientas de selección multicriterio son utilizadas ampliamente en industria para evaluar proveedores, clientes o proyectos, el Proceso Analítico Jerárquico tiene la posibilidad de ser usado en diferentes núcleos de la ciencia sin afectar los resultados (Saaty, 2011, p. 101). Sin embargo, antes de usar esta herramienta se presenta un grupo de herramientas de selección multicriterio.

Herramienta multicriterio	Similitudes con AHP	Diferencias con AHP
El método Macbeth	Se basa en expertos, evalúa cualitativo y cuantitativo	Atributos o características Finitas
Teoría de la utilidad Multiatributo	Tiene alternativas características y criterios como partes de la herramienta.	Se requiere de alto conocimiento por parte del decisor
Relaciones de Superación	Número finito de criterios para hacer comparación uno por uno.	Usan como mecanismo básico comparaciones binarias de alternativas
Proceso analítico en red	Misma base teórica del AHP	Permiten relaciones de interdependencia entre elementos de distintas ramas
Ponderación lineal	Sencillo de utilizar. Aplica pesos a los criterios.	Se emplea cuando se dispone de poca información, en situaciones de incertidumbre

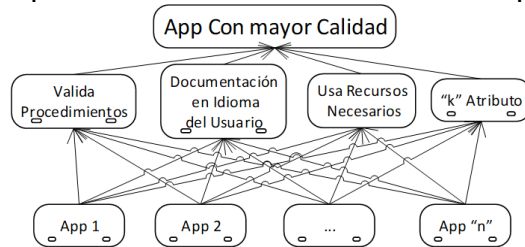
Tabla 1. Similitudes y Diferencias de AHP con otras Herramientas de Análisis Multicriterio (Venegas, 2019 p. 26)

La ventaja de esta herramienta radica en la posibilidad de evaluar “k” alternativas (aplicaciones móviles) y “n” criterios (atributos de calidad). Es posible cuantificar características o criterios tanto cualitativos como cuantitativos, permitiendo una elección neutral de la mejor opción (App móvil médica) para el usuario. El objetivo de la herramienta AHP es elegir la mejor alternativa entre múltiples opciones utilizando como elemento de decisión la comparación por pares. Esta herramienta



se compone de tres niveles; el objetivo, los criterios o características y las alternativas. Las alternativas son las opciones disponibles para elegir, las características son elementos base para tomar la decisión y el objetivo es elegir la alternativa óptima. El proceso utiliza la comparación por pares a partir de una escala de calificación por preferencias donde; 1 = igualmente preferente, 3 = moderadamente preferente, 5 = fuertemente preferente, 7 = muy fuertemente preferente, 9 = extremadamente preferente.

Fig. 1: AHP aplicado a atributos de calidad de aplicación móvil.

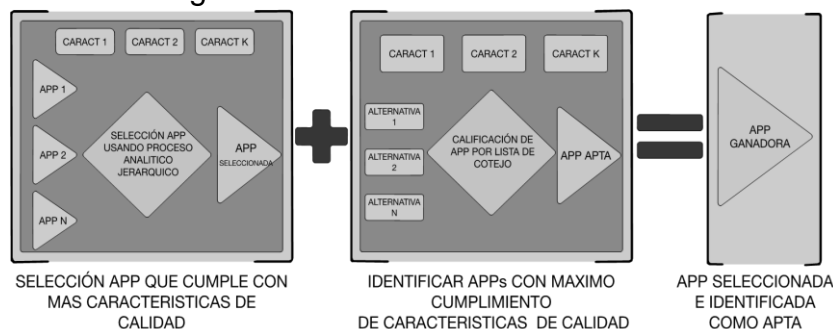


Las calificaciones por pares se realizarán para comparar las características con ellas mismas, posteriormente se comparan las alternativas con sí mismas bajo el enfoque de cada característica.

## 2.1 Aportación

El Proceso Analítico Jerárquico es una buena herramienta para encontrar la mejor alternativa de un grupo de opciones aun cuando todas las alternativas son malas. Por ejemplo, supóngase que se presentan 4 aplicaciones como alternativas y ninguna cuenta con características de calidad aceptables para la rama médica, la herramienta permite seleccionar la mejor propuesta sin importar si todas las propuestas tienen bajo nivel de calidad o son "malas" lo que significa que utilizando el método de selección multicriterio por sí solo es insuficiente para determinar cuál aplicación es recomendable utilizar.

Fig. 2: AHP con Rúbrica de evaluación



Por esta razón se propone una rúbrica de evaluación utilizando la escala de calificación presentada por Thomas Saaty para el AHP, agregando el número cero y contando con 3, 5, 7 y 9 como únicos rangos de calificación. Posterior a la calificación de cada aplicación usando como guía la rúbrica de evaluación en la

Tabla 6, se propone una condición de conteo de valores para validar el máximo cumplimiento de características de calidad. Esta propuesta de métrica no puede determinar la mejor App de forma individual, puesto que cuando un grupo de aplicaciones cumple las mismas condiciones de conteo será calificado con el mismo valor, siendo difícil calificar el grado de calidad de las aplicaciones. La lógica de trabajo se divide en dos partes (ver Figura 2), donde una parte es la selección de la mejor alternativa utilizando el AHP y la otra se encarga de evaluar por separado cada aplicación para validar el cumplimiento de los atributos críticos de calidad. La App móvil óptima se encuentra cuando la aplicación con el puntaje más alto en el Proceso Analítico Jerárquico se encuentra en el rango de aplicaciones aceptadas en la métrica de lista de cotejo. Cuando ninguna App obtiene la calificación adecuada en el sistema de calificación por lista de cotejo, la técnica de decisión multicriterio estará entregando la calificación de las aplicaciones teniendo en cuenta que ninguna es apta para su uso, es decir, el AHP determinaría la mejor aplicación entre un grupo de aplicaciones con calidad no aceptable.

### **3. CASO DE APLICACIÓN: CATEGORIZACIÓN DE CINCO APLICACIONES MÓVILES MÉDICAS QUE MIDEN RÍTMO CARDÍACO**

Para el caso ejemplo se determinaron cinco aplicaciones móviles que miden el ritmo cardíaco. A pesar de tener un funcionamiento similar (todas solicitan poner el dedo índice en la cámara del dispositivo móvil) algunas miden nivel de oxígeno en la sangre e incluso el estado de ánimo. Pero el objetivo principal de las aplicaciones seleccionadas es la determinación del ritmo cardíaco, dichas mediciones se pueden usar de forma general para actividades recreativas, deportivas o para implicaciones médicas, en cualquiera de los casos podría comprometer la salud del usuario y en caso de ser usada para fines médicos la aplicación funcionaría en un equipo móvil como un dispositivo médico. Para calificar de mayor a menor las cinco aplicaciones y determinar cuál de ellas se puede recomendar al usuario para su uso se siguieron dos pasos:

Paso 1 Determinación de elementos de calidad críticos para una aplicación móvil del área médica: Las características de calidad de software propuestas por los diferentes organismos basan sus propuestas en la norma ISO/IEC 9126 desarrollada para identificar los atributos de calidad para un software y de las cuales se extraen los principales atributos de calidad y en la ISO/IEC 14598 la cual da una descripción de los procesos para evaluar software y guías para evaluar la calidad de los productos. Según los atributos de calidad presentes en la ISO/IEC 9126 para los productos software de la cual se basan las agencias reguladoras para hacer recomendaciones de calidad para aplicaciones móviles para la salud se utiliza una muestra de atributos para considerar una aplicación médica móvil como confiable. Se debe aclarar que la propuesta puede evaluar “k” características de calidad, sin embargo, se presentan nueve características a modo de muestra para poder probar la herramienta evaluadora.

	<b>Elementos de calidad Críticos</b>	<b>¿Quién propone?</b>
a	Se validan todos los procedimientos de registros electrónicos y firmas electrónicas.	Titulo 21 del código de regulaciones federales de los Estados Unidos, Parte 11 del 21 CFR (11.10 - 11.30 - 11.50 - 11.200 - 11.300)
b	Documentación completamente en idioma del usuario comprensible y suficiente sobre uso de la aplicación, fuentes de información y funcionamiento de la aplicación.	Propuesto a partir del análisis de características críticas del presente artículo (Paso 1 de esta sección)
c	Utiliza el dispositivo móvil como un equipo médico regulado (Mide, monitorea y/o proporciona datos de salud)	Documento número 1741 de la Agencia de drogas y alimentos de los Estados Unidos (FDA) apéndice C
d	Solicita y utiliza únicamente los recursos necesarios para cumplir con el objetivo del diseño del producto.	Propuesto a partir del análisis de características críticas
e	Es plenamente identificado el creador de la aplicación y o responsables del producto y se puede tener contacto con ellos	Agencia de Calidad Sanitaria de Andalucía
f	“La App de salud advierte de aquellas actualizaciones que inciden o modifican funcionamientos o contenidos sobre salud o cualquier otro dato sensible.”	Agencia de Calidad Sanitaria de Andalucía
g	“Se identifican los riesgos que el manejo de la App de salud puede suponer para la seguridad del paciente.” y toma las acciones preventivas respectivas.	Agencia de Calidad Sanitaria de Andalucía
h	“La App de salud contiene la fecha de la última revisión realizada sobre el material publicado.”	Agencia de Calidad Sanitaria de Andalucía
i	“La App de salud dispone de un sistema de ayuda sobre su manejo.” amplio y suficiente por medio escrito.	Agencia de Calidad Sanitaria de Andalucía

Tabla 2. Elementos de Calidad para Aplicaciones Móviles (FDA, 2016)

Paso 2 Uso de rúbrica de evaluación y herramienta proceso analítico jerárquico para clasificar aplicaciones médicas: La propuesta de evaluación de aplicaciones móviles médicas se realizará en dos etapas: en la primera se utiliza el Proceso Analítico Jerárquico para calificar las aplicaciones móviles médicas. En la segunda se evaluará por medio de la lista de cotejo los atributos de calidad presentes en la aplicación, como resultado se genera un dictamen de calificación. Al finalizar las

dos etapas se obtendrán; la aplicación que más cubre las características establecidas según los expertos y las aplicaciones recomendadas por su calidad funcional.

### 3.1 Aplicando Proceso Analítico Jerárquico

El objetivo es seleccionar entre un grupo de opciones la aplicación móvil que cumpla de mejor manera los atributos de calidad. El primer paso de la metodología consiste en hacer una evaluación por pares de los nueve atributos de calidad. En la Tabla 3 se evalúa la importancia de cada característica con las demás, siendo la calificación 1 en la diagonal principal debido a que se compara una característica con ella misma (Escala de preferencias establecida por Thomas Saaty). En el procedimiento una vez se tiene la comparación por pares por parte de los expertos Tabla 3 se suman las columnas y se divide cada elemento por el total de cada columna los resultados.

Características de Calidad									
Características	a	b	c	d	e	f	g	h	i
a	1	9	1	9	9	9	7	7	5
b	1/9	1	1/9	3	3	5	3	3	5
c	1	9	1	9	9	5	5	9	7
d	1/9	1/3	1/9	1	7	3	3	9	5
e	1/9	1/3	1/9	1/7	1	3	5	9	7
f	1/9	1/5	1/5	1/3	1/3	1	3	9	9
g	1/7	1/3	1/5	1/3	1/5	1/3	1	9	9
h	1/7	1/3	1/9	1/9	1/9	1/9	1/9	1	3
i	1/5	1/5	1/7	1/5	1/7	1/9	1/9	1/3	1

Tabla 3. Comparación por Pares de Características

En la Tabla 5 se observa que las características más importantes son la a) con un 29% y la característica c) con un 27,45%, por otra parte, la aplicación menos importante según la opinión de los expertos es la i) con una calificación de 2%. Las características fueron evaluadas, pero es necesario corroborar que la tasa de consistencia (CR por sus siglas en inglés) de la matriz no supere el 10%. Esto quiere decir, que las evaluaciones de los expertos para la matriz inicial de la Tabla 3 son coherentes. Para poder calcular la tasa de consistencia CR en Figura 3 es necesario encontrar el índice de consistencia el cual se representa con CI, por sus siglas en inglés (Saaty, 2011, p180)

Fig. 3: Calculo de Tasa de Consistencia o CR

$$CI = (\lambda_{\max} - n)/(n - 1) \quad (1)$$

$$CR = CI/RI < 0,1 \sim 10\% \quad (2)$$

$$CR = 13,78074/1,45 = 9,5039 \quad (3)$$

El cálculo de  $\lambda_{\max}$  corresponde a la multiplicación de los valores resultado por el total de la suma de cada columna. Es necesario utilizar también una tabla de índices de consistencia aleatoria o RI que se muestra en la Tabla 4.

N	1	2	3	4	5	6	7	8	9	10
RI	0	0	0,58	0,9	1,12	1,24	1,32	1,41	1,45	1,49

Tabla 4. Índices de Consistencia Aleatoria

La tasa de consistencia CR es de 9,50% (ver Figura 3), por lo cual se puede considerar que la matriz inicial aportada por los expertos al evaluar las nueve características es consistente, validando los resultados de esta primera etapa. El siguiente paso es evaluar por pares a las alternativas respecto de cada característica, el procedimiento algebraico es el mismo con lo cual se entrega en la Tabla 5 los resultados de dichos análisis.

Tabla de Resultados (Calificación de App Médicas)										
APP	a	b	c	d	e	f	g	h	i	Calificación
App 1	0,2	0,1	0,27	0,09	0,08	0,07	0,06	0,02	0,02	15,3
App 2	20	3,60	20	4,36	7,32	20	16,5	6,31	7,69	25,1
App 3	20	11,4	20	46,4	56,5	20	16,5	44,2	23,08	17,8
App 4	20	28,3	20	8,88	6,99	20	6,31	16,5	23,08	22,1
App 5	20	28,3	20	20,1	18,6	20	44,1	16,5	23,08	19,5

Tabla 5. Resultados Luego de Aplicar AHP

Se puede ver en la Tabla 5 que la aplicación 2 es la mejor calificada con 25,05% y el producto que según la evaluación de los expertos tiene menores características de calidad es la App 1 con un 15,29% de calificación. Sin embargo, la evaluación del AHP califica las aplicaciones sin tener en cuenta cuando estos productos son aptos para el uso sin poner en peligro la salud de las personas. Para realizar esta prueba se debe hacer una evaluación por lista de cotejo.

### 3.2 Aplicando Rúbrica de Evaluación

En esta etapa se propone evaluar las aplicaciones que se presentan para determinar en qué casos existe riesgo para la salud al utilizarlas. Para tal fin se presenta una Rúbrica de Evaluación para evaluar las características deseables, en la Tabla 6 se observan los primeros cuatro atributos de calidad y las opciones de calificación según el caso. Una aplicación que no cumple el atributo de calidad obtiene una calificación de cero "0", en cambio, cuando el atributo es cumplido a cabalidad su calificación máxima será de "9". El rango de calificaciones propuesto se basa en los parámetros presentados por Thomas L. Saaty para el AHP.

<b>Rúbrica de calificación de características de calidad</b>					
	<b>0</b>	<b>3</b>	<b>5</b>	<b>7</b>	<b>9</b>
Aplicación validada por parte de un grupo de especialistas médicos para su publicación.	No hay evidencia de validación por parte de expertos.	Existe duda razonable sobre la autenticidad de la validación médica para la publicación de la App.	Esta validada la aplicación, pero no por expertos del área.	Aplicación validada por expertos médicos de otro país.	Aplicación validada por expertos médicos en el país de publicación.
Documentación completamente en idioma oficial. Se enuncian las fuentes de información y funcionamiento de la aplicación.	No hay documentación de ningún tipo.	Existen dudas sobre la exactitud de la traducción y sobre las fuentes de información.	Documentación parcialmente en el idioma oficial o mal traducido, pero se entiende la documentación sobre el funcionamiento del producto.	Traducción completa en el idioma oficial. Se presentan fuentes de información. Sin embargo, no se explica el funcionamiento del producto.	Documentación en el idioma oficial. Fuentes de información actuales y verídicas. Funcionamiento de la aplicación claro y preciso.
Todas las decisiones y diagnósticos validados por expertos médicos.	Toma decisiones y genera diagnósticos sin validación médica o no toma decisiones.	El producto toma decisiones y genera diagnósticos con validaciones de personal no experto.	Toma decisiones y genera Diagnósticos validados por personal médico en otro idioma.	Genera Diagnósticos validados por personal médico indirectamente.	Genera diagnósticos validados por personal médico en tiempo real y de forma directa.

Solicita y utiliza únicamente los recursos necesarios para cumplir con el objetivo del diseño del producto.	Solicita y/o utiliza recursos del dispositivo móvil más allá de las necesidades básicas de la App, accede a información privada del usuario.	Usa recursos más allá de las necesidades de la App, pero no accede a la información privada del usuario.	Solicita recursos adicionales al objetivo de funcionamiento de la App, pero el usuario puede negarse al acceso y usar la aplicación.	No solicita ni utiliza recursos adicionales al objetivo de funcionamiento de la App, pero podría poner en peligro la privacidad de los datos del usuario.	Utiliza recursos necesarios para mantener calidad funcional del producto, se respeta privacidad del usuario.
---	--	--	--	---	--

Tabla 6. Rúbrica de Evaluación para Características de Calidad en app (ISO, 2013, 2015 y 2016)

#### 4. RESULTADOS

Una vez registradas las calificaciones se realiza una revisión de valores en la Tabla 7 para cada columna utilizando como referencia las condiciones de la Tabla 8 se asigna la calificación obtenida. Las condiciones para cada columna indican cuando una App tiene la mitad de las características con calificaciones de 7 o 9 entonces su calificación corresponderá a 5. Cuando una app tiene el 80% de las calificaciones con 7 o 9 su calificación será de 7, cuando la app cuenta con el total de las características de calidad con calificación de 7 o 9 tendrá una calificación de 9. El resultado del análisis para cada columna se registra en la Tabla 7 en la parte de “calificación obtenida”.

<b>Características (elementos de calidad)</b>	<b>APP 1</b>	<b>APP 2</b>	<b>APP 3</b>	<b>APP 4</b>	<b>APP 5</b>
a	0	0	0	0	0
b	3	7	9	9	9
c	0	0	0	0	0
d	3	9	5	7	7
e	3	9	3	5	3
f	0	0	0	0	0
g	3	3	0	5	3
h	0	5	3	3	3
i	5	9	9	9	9
Calificación obtenida	0	0	0	0	0
Usar tabla de validación para reportar calificaciones					

Tabla 7. Calificaciones de APP

La calificación por lista de cotejo para las aplicaciones móviles sirve para detectar a partir de una validación de calidad funcional cuales productos son confiables y pueden ser usados sin temer un impacto en la salud.

<b>Validación de la app</b>	
<b>Si la aplicación posee</b>	<b>Califica</b>
Menos de la mitad de las características con 7 o 9	0
La mitad de las características con 7 o 9	5
8 de cada 10 características con 7 o 9	7
todas las características tienen 7 o 9	9

Tabla 8. Métrica de Evaluación para Lista de Cotejo

Según el resultado mostrado en la Tabla 7, ninguna de las cinco aplicaciones analizadas tiene una calificación aprobatoria, lo cual significa que ninguna es una aplicación confiable, puesto que no cumplen con los atributos de calidad de la norma ISO/IEC 9126 en los atributos de calidad muestra utilizados para la evaluación de las aplicaciones.

## 5. CONCLUSIONES

Combinar el proceso analítico jerárquico con una escala de calificación permite determinar si las características de un producto software cumplen con una regulación determinada, no necesariamente aplicado a productos médicos, pero si aplicado a productos regulatorios. Aun así, puede utilizarse para tener un valor duro de calificación sobre algún aspecto determinado, por ejemplo, utilizando como experto evaluador a los clientes, usuarios o administradores del sistema, permitiendo que estos evalúen y extraigan una calificación sobre el producto que les presentan. En definitiva, lo que se busca es tener una forma de evaluar software sin tener que usar métodos subjetivos convirtiendo valoraciones cualitativas en cuantitativas.



## REFERENCIAS

Fernandez, S., Ramos, J. (2014). New technologies, apps and their application in evidence-based clinical practice. *Enferm Clin* núm 24. doi: <https://doi.org/10.1016/j.enfcli.2014.02.001>, Recuperado en septiembre de 2018. pp. 99-101

Garcia, S., Martinez, J. (2016). Apps de Salud: Nuevas Herramientas para el cuidado del Paciente Cardiológico. *Enfermería en Cardiología* núm 22. Pp 28-34

International Organization for Standardization. (2013). ISO 15504: Determinación de la Capacidad de Mejora del Proceso de Software. Ginebra: ISO

International Organization for Standardization. (2015). ISO 15288: Ingeniería de sistemas y software. Procesos del ciclo de vida del sistema. Ginebra: ISO

International Organization for Standardization. (2016). ISO 13485: Sistemas de Gestión de la Calidad de Equipos Médicos - Guidance. Ginebra: ISO

Muñoz, C., Piattini, G., Moraga, A. (2010). Calidad del producto y proceso software. Madrid: Ra-Ma. ISBN: 8478979611,9788478979615.

Redrován, F., Loja, N., Correa, k., Piña, J. (2017). Comparación de métricas de calidad para el desarrollo de aplicaciones web. *3C Tecnología. Glosas de Innovación aplicadas a la pyme* núm 7(3). <http://dx.doi.org/10.17993/3ctecno.2018.v7n3e27.94-113/>, Recuperado en Septiembre de 2018. pp. 94-113

Saaty, T.L. (2011). *Fundamentals of Decision Making and Priority Theory with AHP*. RWS Publications, Pittsburg. ISBN: 0962031763,978-0962031762

Tuya, J., Ramos, I., Dolado, J. (2007). *Técnicas cuantitativas para la gestión de ingeniería de software*. Oleiros: Netbiblo. ISBN:9788497452045.

U.S. Food and Drug Administration. (2016). *Codigó de las Regulaciones Federales 21 parte 11: Electronic Records; Electronic signatures – Guidance*. Estados Unidos: GMP publications. ISBN: 1933734035

U.S. Food and Drug Administration. (2016). Codigó de las Regulaciones Federales 21 parte 820: Quality System Regulation– Guidance. Estados Unidos: GMP publications. ISBN: 1933734507

Venegas, J., (2019). *Modelo de proceso de desarrollo de software MW para productos regulados*. Tesis de doctorado, Universidad Juárez Autónoma de Tabasco, México.

## NOTAS BIOGRÁFICAS



**M.C.I. Jonathan Venegas Sandoval** Estudiante de Doctorado en Ciencias de la Computación por la Universidad Juárez Autónoma de Tabasco, Maestro en Ciencias de Ingeniería (2014) por la Universidad Politécnica de Aguascalientes. Ingeniero Industrial (2008) Universidad Autónoma de Colombia.



**Dr. Francisco Javier Álvarez Rodríguez**, Profesor de Ingeniería de Software adscrito al Departamento de Ciencias de la Computación, Universidad Autónoma de Aguascalientes (U.A.A.). Doctor en Metodología de la Enseñanza por el IMEP (México). Doctor en Ingeniería por la UNAM (México). Ha sido Decano del Centro de Ciencias Básicas en la U.A.A., así como Jefe de Departamento de Sistemas Electrónicos. Miembro de núcleos académicos de diversos posgrados de la U.A.A.



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 2.5 México.

*Recibido 20/07/2019*

*ReCIBE, Año 8 No. 2, Noviembre 2019*

*Aceptado 04/09/2019*

## **Arquitectura de un juego serio inteligente basado en retos de matemáticas básicas**

**Carlos Enrique Flores Méndez<sup>1</sup>**  
cfloresm932@alumno.uaemex.mx

**Maricela Quintana López<sup>1</sup>**  
mqintanal@uaemex.mx

**Héctor Rafael Orozco Aguirre<sup>1</sup>**  
hrorozcoa@uaemex.mx

**Ivonne Rodríguez Pérez<sup>1</sup>**  
irodriguezp@uaemex.mx

<sup>1</sup>Centro Universitario UAEM Valle de México  
Universidad Autónoma del Estado de México

**Resumen:** En México, en el mes de junio de 2018, el Instituto Nacional para la Evaluación de la Educación aplicó a más de 100,000 estudiantes de sexto año de primaria, la prueba PLANEA, la cual se encarga de evaluar aprendizajes de lenguaje, comunicación y matemáticas. Los resultados muestran que, en matemáticas, el 59% de los estudiantes se encuentran en un nivel insuficiente. Para solventar estas deficiencias durante los últimos años se han desarrollado nuevas técnicas de aprendizaje, entre ellas, el aprendizaje basado en retos y los juegos serios, los cuales, tienen como fundamento involucrar a los estudiantes en un entorno de aprendizaje en un contexto de entretenimiento. Nuestra propuesta es crear un juego serio que incorpore estas técnicas, con el objetivo de que los estudiantes adquieran o refuercen sus conocimientos en matemáticas básicas. Como un primer paso, una arquitectura para el juego serio se propone y se presenta en este trabajo.

**Palabras clave:** Juegos serios, aprendizaje basado en retos, arquitectura de juegos serios, inteligencia artificial.

**Abstract:** On June 2018, the National Institute for the Evaluation of Education in Mexico, applied the PLANEA test to roughly 100,000 students in the sixth year of elementary school, that evaluates language, communication and mathematics learning. The results show that, in mathematics, 59% students achieved insufficient level. To solve these deficiencies in recent years, new learning techniques have been developed, including challenge-based learning and serious games, which are based on involving students in a learning environment in a context of entertainment. Our proposal is to create a serious game that incorporates these techniques, with the intend of acquire or reinforce their basic mathematic knowledge. An architecture for the serious game is proposed and presented in this paper.

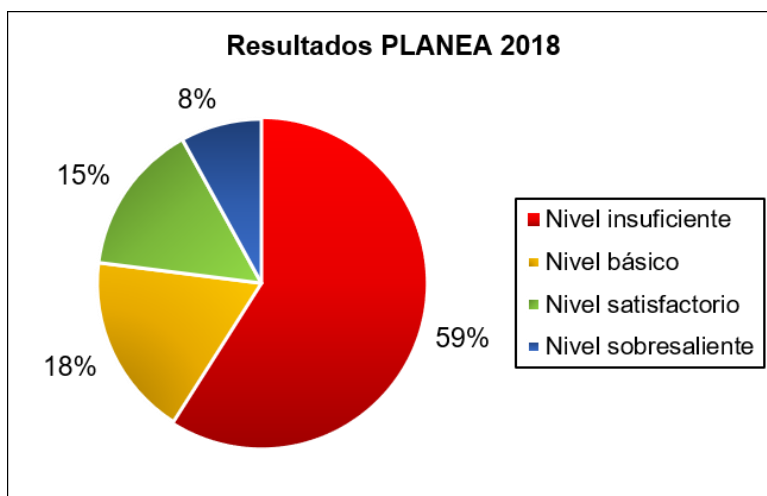
**Keywords:** Serious games, challenge-based learning, serious game architecture, artificial intelligence.

## 1. Introducción

En México, el Instituto Nacional para la Evaluación de la Educación (INEE) es un organismo público autónomo descentralizado de la Secretaría de Educación Pública (SEP). De acuerdo con su reporte: PLANEA – Resultados Nacionales (INEE, 2018), en el mes junio del año 2018, este organismo, en conjunto con la SEP, aplicó en 3,753 escuelas del país y a 104,973 alumnos de sexto año de primaria la prueba PLANEA, la cual consta de 141 reactivos para evaluar aprendizajes de Lenguaje y Comunicación y 147 que evalúan conocimientos de Matemáticas.

Los resultados obtenidos y presentados en el reporte de noviembre de 2018 indican que en Matemáticas el 59% de los estudiantes tienen un nivel insuficiente, 18% un nivel básico, 15% se halla en nivel satisfactorio y, por último, un 8% se localiza en nivel sobresaliente, en la figura 1 se muestra la distribución de resultados de esta prueba.

Desde el año 2005 en México se aplican pruebas estandarizadas para medir el desempeño de los estudiantes, tal es el caso de la prueba Enlace y su sucesor: PLANEA, los propósitos de esta prueba pretenden aportar información útil para que tanto autoridades, consejos técnicos y profesores, mejoren los procesos educativos y sus prácticas de enseñanza (Barragán, 2017). Otro de los objetivos de la prueba PLANEA es informar cuáles son las competencias en las que los estudiantes están fallando, sin embargo, esta prueba no está diseñada para advertir el motivo de estas deficiencias (Santibáñez, 2015). A pesar de los esfuerzos gubernamentales, los resultados indican que una porción significativa de los estudiantes mexicanos no está aprendiendo lo suficiente, y lo más preocupante es, que el desempeño de los estudiantes ha ido en declive (Santibáñez, 2015).



**Figura 1.** Distribución de resultados de la prueba PLANEA

En la sección 2 de este trabajo, se tratará del aprendizaje basado en retos, mientras que en la sección 3 se hablará de los juegos serios, así como de la arquitectura general de estos. En la sección 4, se presenta el juego serio y la arquitectura propuesta. Finalmente, en la sección 5 se presentan las conclusiones y el trabajo futuro.

## **2. Aprendizaje basado en retos**

La tecnología educativa está basada en diversas teorías de aprendizaje y enfoques pedagógicos como lo es el constructivismo (Medina, 2007), en este sentido (Piaget, 1978) menciona que el conocimiento se edifica a través de la experiencia, donde cada persona construye significados a medida que se involucra en la realización de actividades, de esta manera, el sujeto realiza esquemas mentales que en un principio no se entienden, ni se utilizan de manera inmediata, sino que con cada experiencia se va aprendiendo.

Durante los últimos años se han creado nuevas técnicas de aprendizaje entre las cuales se encuentra el Aprendizaje Basado en Retos, el cual de acuerdo con el Instituto Tecnológico de Monterrey en su Reporte Edutrends (2015), el aprendizaje basado en retos es un enfoque pedagógico que involucra activamente al estudiante en una situación problemática real. Esta situación debe ser relevante y de vinculación en un entorno que sirva para definir un reto y que a los estudiantes les permita implementar una solución posible.

El aprendizaje basado en retos se basa en el aprendizaje vivencial, el cual tiene como fundamento principal que los estudiantes aprenden mejor cuando participan en experiencias de aprendizaje abiertas, en lugar de cuando se participa en forma pasiva y secuencial, por otro lado, se han cuantificado numerosas ventajas de este enfoque:

- Los estudiantes logran una comprensión más profunda de los temas, además, aprenden a diagnosticar y definir los problemas antes de proponer soluciones (ITM, 2015).
- Los estudiantes se involucran tanto en la definición del problema, como en la solución a desarrollar (Johnson, 2015)
- Los estudiantes fortalecen la conexión entre lo que aprenden en la escuela y lo que perciben en el mundo que los rodea (Johnson, 2015).

Una de las características fundamentales del Aprendizaje Basado en Retos está relacionado con la evaluación, si bien, este enfoque no cuenta un método general y unificado, es posible identificar estrategias para valorar los procesos y las soluciones presentadas para proveer a los estudiantes de una retroalimentación de cada una de las fases a las que se ha sometido. De forma general, la evaluación de los estudiantes debe considerar dos tipos de estrategias de evaluación: la formativa y la sumativa, la formativa ocurre durante todo el proceso, guiando y facilitando el

aprendizaje; mientras que la sumativa valora el progreso en puntos clave o la finalización de este (ITM, 2015).

### 3. Juegos serios

La tecnología educativa incorpora a las Tecnologías de la Información y Comunicación (TIC) como apoyo a los procesos de aprendizaje, de acuerdo con el Glosario de Términos de Tecnología Educativa (UNESCO, 1984), la tecnología educativa es un modo sistemático de concebir, aplicar y evaluar los procesos de enseñanza y aprendizaje, reconociendo los recursos técnicos, humanos y sus interacciones.

En este sentido, dentro de los múltiples servicios que ofrecen las TIC, se encuentran los servicios educativos como lo es el e-learning, el cual es un tipo de enseñanza, que se caracteriza por la utilización de Internet como canal de distribución de conocimiento a través de distintos dispositivos, esto, mediante sistemas informáticos como son: el software educativo, los sistemas tutores inteligentes y los juegos serios.

Un juego serio constituye una herramienta que tiene como objetivo principal el aprendizaje y donde los jugadores se enfrentan a retos individuales o colectivos con el objetivo de entretener a los participantes, mientras estos, ponen en práctica sus habilidades (Urquidí, Cristina, & Aznar, 2015). Los juegos serios aportan ventajas para los estudiantes: Marcano (2008), explica que, a nivel sensorial, los juegos serios permiten mejorar la agudeza de la percepción visual, la rapidez de reflejos, la memoria y la capacidad de seguir instrucciones; a nivel cognitivo, facilitan el desarrollo de estrategias para la solución de problemas, es decir, aprenden con las nuevas experiencias que se les presentan y aprenden a desarrollar estructuras de pensamiento.

La clasificación de los juegos serios no es un tema que este propiamente consolidado, sin embargo, se pueden definir dos tipos: los orientados a procesos y los orientados a resultados. Cada tipo tiene su propia clasificación, estas se presentan en las tablas 1 y 2.

<b>Clasificación</b>	<b>Descripción</b>
<b>Descubrimiento</b>	Proporcionan nuevas perspectivas en temas de marketing, equipo de trabajo, clientes y votantes.
<b>Toma de decisiones</b>	Proporcionan mecanismos para que la toma de decisiones sea más rápida.
<b>Simulación</b>	Son simulaciones cara a cara o virtuales de situaciones que podrían presentarse en el mundo real

**Tabla 1.** Clasificación de los juegos serios orientados a procesos.

<b>Clasificación</b>	<b>Descripción</b>
<b>Educación</b>	Ejercicios que comunican y evalúan habilidades, conceptos y capacidades.
<b>Toma de decisiones</b>	Proporcionan experiencias que construyen marcas que influyen en formar una opinión.
<b>Simulación</b>	Proporcionan insignias puntajes y otras recompensas basadas en el trabajo.

**Tabla 2.** Clasificación de los juegos serios orientados a resultados

El juego serio que se está desarrollando para el aprendizaje de las matemáticas, pertenece a los orientados a resultados, y específicamente, a la clase educacional. En este trabajo, se presenta una propuesta de su arquitectura que permita cumplir su propósito.

Por otro lado, la aplicación de juegos serios, no importando su orientación y clasificación, se considera beneficiosa y estos se han utilizado en diferentes tipos de ambientes por ejemplo el militar, el político, en entornos empresariales y corporativos y por supuesto en el área educativa (Marcano, 2008). Por ejemplo, en el sector médico se encuentra el juego serio “Pulse”, el cual reproduce las condiciones de una sala de emergencias de un hospital, donde el objetivo central, es que los futuros enfermeros puedan practicar todo lo aprendido en las clases teóricas, identificando los problemas de cada paciente y aprendiendo a priorizar a los más graves y en base a ello aplicar las medidas apropiadas para no dejar que mueran (GameLearn, 2017). Otro ejemplo, es “Dragon Box Elements”, en el cual los jugadores tienen que construir un ejército para derrotar un dragón, para esto los jugadores tienen que aprender las bases de la geometría (GameLearn, 2017).

### **3.1 Arquitectura de un juego serio**

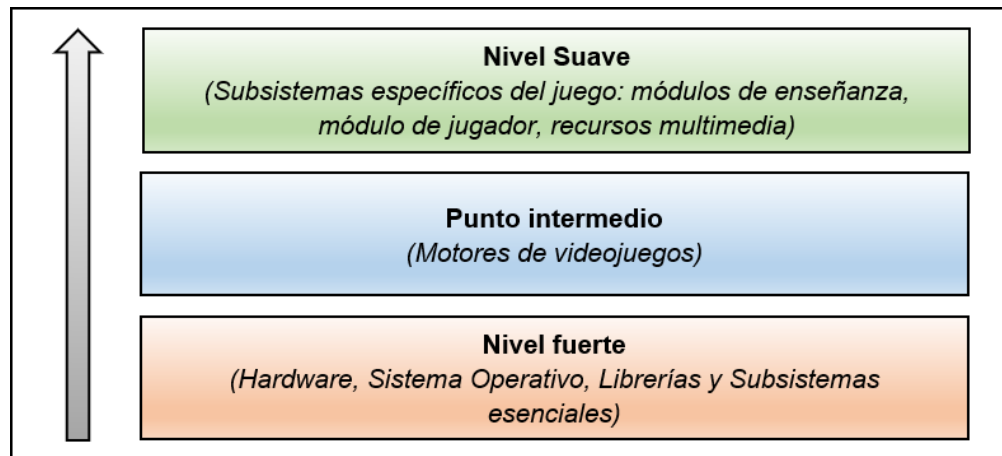
El diseño de un juego serio requiere de un fuerte conocimiento que involucra diversas disciplinas como: la programación, diseño de componentes de juego, dominio del aprendizaje y elementos pedagógicos, el objetivo del diseño de un juego serio es tener una combinación de aprendizaje y diversión, sin embargo el diseño de un juego serio sigue siendo difícil de lograr (Walid Mestadi, 2018)

Los juegos serios están usualmente basados en la simulación de un contenido de aprendizaje que se debe representar; desde el punto de vista del diseño de un juego serio, el diseño puede lograrse por medio de una aproximación global, llamada aproximación endógena, la cual consiste en que el contenido de aprendizaje esté fuertemente ligado con la estructura general del juego y sus reglas. Para el diseño de juegos serios, se considera que la aproximación endógena es un buen enfoque, ya que la jugabilidad está integrada con el contenido de aprendizaje de una manera natural (Walid Mestadi, 2018), además de tener una importancia contextual, ya que el jugador se encuentra en el centro del proceso de desarrollo, por lo que el juego serio puede contar con objetivos de aprendizaje más complejos (Fullerton, 2008).



En esta aproximación, no se especifica cómo debe representarse el contenido de aprendizaje dentro del juego, sin embargo, existen elementos comunes que debe tener: los objetivos pedagógicos, la simulación, las interacciones de los jugadores con la simulación, los problemas a los que debe enfrentarse el jugador y el progreso que lleva en la solución de estos.

De forma general, la arquitectura de un juego serio con una aproximación endógena se puede definir en dos niveles o arquitecturas: la fuerte y la suave. El nivel o arquitectura fuerte es aquella que concentra el hardware y un sistema operativo para que los diseñadores, mediante un motor de juegos, creen un juego concreto. A este nivel también se le conoce como la arquitectura del sistema (Capdevilla, 2013). Por otro lado, el nivel o arquitectura suave, comprende la creación del juego concreto utilizando los módulos que los diseñadores del juego consideren necesarios. En la Figura 2, se puede observar la arquitectura general de un juego serio.



**Figura 2.** Arquitectura general de un juego serio

Esta no es la única propuesta de una arquitectura de un juego serio, por ejemplo, (Ricardo Emmanuel Gutiérrez-Hernandez, 2013) propone una arquitectura con aspectos culturales para un juego serio multicultural que cuente con las características de diversas actividades lúdicas y aprendizaje.

A continuación, se presentan las ideas básicas del juego y la arquitectura propuesta para alcanzar los objetivos.

#### **4. Juego serio basado en retos y la arquitectura propuesta**

El juego serio tiene como objetivo que los estudiantes adquieran o refuercen sus conocimientos en matemáticas básicas, es un juego de tipo educativo, en el que el jugador toma el rol de un navegante que se vio envuelto en una extraña tormenta que lo hizo perder el conocimiento y su nave encalla en una isla. Para salir de esta situación, el navegante debe encontrar las 6 piezas faltantes de su embarcación, para ello, el jugador se enfrentará a diversos retos que exponen desafíos matemáticos específicos que van acorde a las competencias evaluadas en la prueba PLANEA.

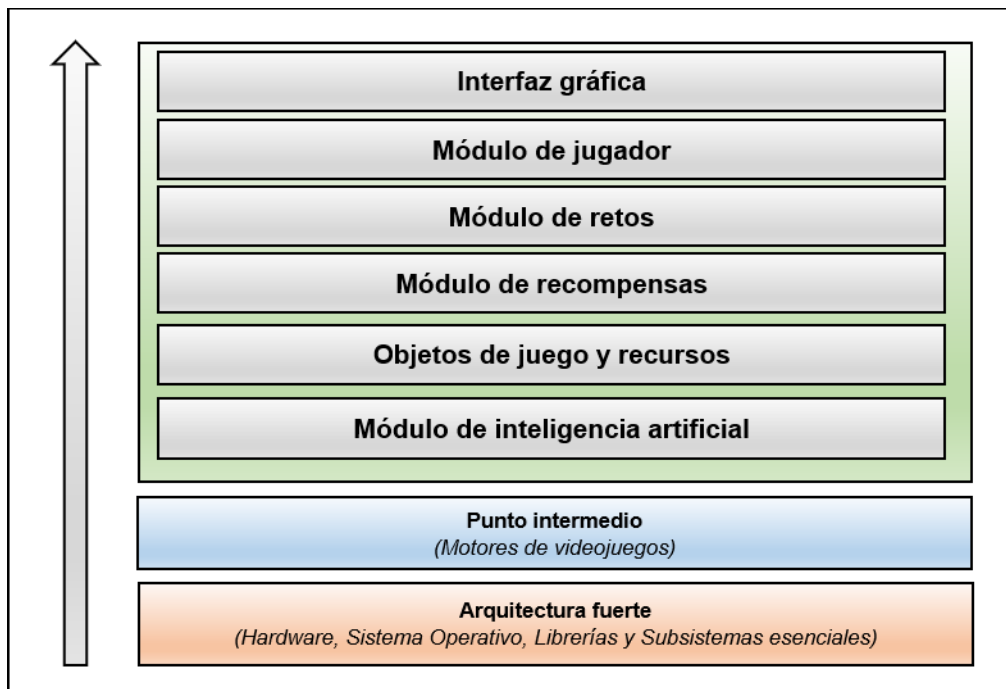
Para el juego, se propone una arquitectura suave con aproximación endógena, donde cada uno de los módulos estén completamente integrados para cumplir con el objetivo de incorporar la técnica de aprendizaje basado en retos. En la figura 3, se muestra la arquitectura general propuesta del juego serio y en la figura 4 cómo sus elementos se relacionan.

A continuación, se explica cada uno de los módulos:

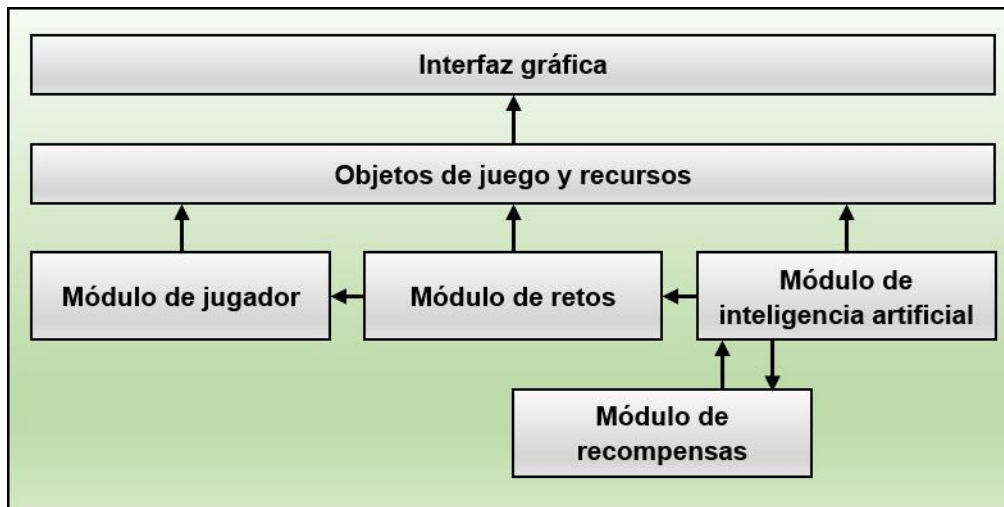
**Interfaz gráfica:** es la representación visual de todos los elementos gráficos y auditivos del juego, que en conjunto proporcionan una experiencia gráfica e inmersiva con la que el jugador interactúa a través de la historia resolviendo diferentes retos.

**Módulo del jugador:** este módulo se encargará de representar al jugador dentro del juego, así como de almacenar los atributos del jugador, como lo son su nombre, el avance y el desempeño que ha tenido durante el juego.

**Módulo de retos:** se encarga de proveer al jugador los retos basados en las competencias matemáticas de la prueba PLANEA. Los retos se presentan en forma de misiones, en donde el jugador se enfrenta a una problemática que tiene múltiples soluciones, estas misiones implican un desafío que le permitirá al jugador generar ideas para probar su destreza y presentar una solución.



**Figura 3.** Arquitectura propuesta para un juego serio basado en retos.



**Figura 4.** Interrelación de los módulos de la arquitectura propuesta

**Módulo de recompensas:** este módulo implementa la evaluación del reto de forma sumativa; se encarga de evaluar la solución presentada por el jugador al final de cada reto, para finalmente proporcionar una recompensa y una retroalimentación.

**Objetos y recursos del juego:** se encarga de almacenar todos los elementos pertenecientes al juego en su forma lógica, como lo es el jugador, los enemigos, los escenarios y los elementos necesarios que conforman los retos.

**Módulo de inteligencia artificial:** Este módulo se encarga de varias tareas, una de ellas es implementar la evaluación formativa del estudiante proporcionando, durante todo el reto, retroalimentación que facilite el aprendizaje. Por otro lado, dependiendo de la evaluación realizada por el módulo de recompensas, permitirá aumentar o disminuir la dificultad del juego.

La forma en que esta arquitectura será implementada provee un escenario de libre exploración que permitirá a los jugadores involucrarse en la resolución del reto, durante el progreso de este juego serio el jugador se someterá a: la fase de jugabilidad que se muestra en el diagrama de secuencia de la figura 5 y la fase de aprendizaje que se muestra en la tabla 3.

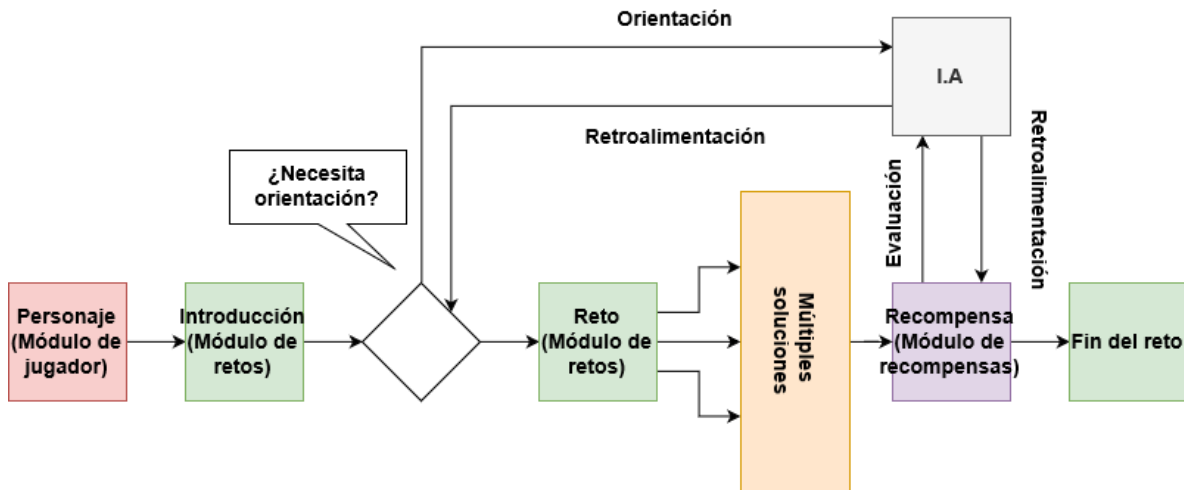


Figura 5. Diagrama de secuencia de la arquitectura propuesta.

Módulo	Fase de jugabilidad	Fase de aprendizaje
<b>Módulo de jugador</b>	El jugador configura al personaje que controlará en el juego	Aprende a configurar el juego
<b>Módulo de retos</b>	Recibe una introducción de la situación específica del nivel del juego y sus objetivos. El jugador intenta resolver el reto propuesto.	Aprende las mecánicas del juego y cómo realizar la interacción específica del nivel. El jugador se somete a la experimentación de soluciones dentro del reto
<b>Módulo de recompensas</b>	Se evalúa el desempeño del jugador y se determina la recompensa apropiada	El jugador aprende de sus errores en caso de fallo para obtener una mejor recompensa
<b>Módulo de inteligencia artificial</b>	El jugador obtiene información valiosa y consejos de cómo pasar el nivel del juego dependiendo de su desempeño. Además, incrementa o disminuye la dificultad del juego	El jugador recuerda los consejos del avatar y los aplica al jugar de nuevo el nivel

Tabla 3. Fases de progresión del juego serio basado en la arquitectura propuesta.

## 5. Conclusiones y trabajo futuro

La arquitectura suave propuesta permite la inclusión de los módulos de retos, de recompensas, de inteligencia artificial, y del jugador, los cuales al interactuar entre sí y con los objetos y recursos del juego permiten una experiencia diferente para cada jugador de acuerdo con sus habilidades, conocimientos.

Particularmente, el uso de un módulo de inteligencia artificial que permita variar la dificultad de un nivel de acuerdo con el desempeño del jugador, en conjunto con las técnicas de aprendizaje vivencial como lo es el aprendizaje basado en retos, dentro

de la arquitectura del juego serio permite aprovechar el potencial y objetivo general de este tipo de software: entretener mientras se enseña.

Como trabajo futuro, se encuentra implementar esta arquitectura para crear un juego concreto que contenga las características apropiadas para implementar el aprendizaje basado en retos, y que sirva como herramienta adecuada para que los estudiantes mexicanos que cursan el sexto año de educación básica obtengan un aprendizaje significativo y en consecuencia mejoren sus habilidades matemáticas que son evaluadas en la prueba PLANEA.

### **Agradecimientos**

- Agradezco a la Universidad Autónoma del Estado de México por otorgarme la beca de estudios avanzados y por apoyarme con las herramientas necesarias para realizar mi maestría.
- Agradezco al Consejo Nacional de Ciencia y Tecnología CONACYT por la asignación de la beca que me ha permitido estudiar mi Maestría en Ciencias de Computación.
- Un agradecimiento especial a mis tutores por su guía y valiosos consejos que me han permitido realizar este trabajo.

### **REFERENCIAS**

- Barragán, A. S. (17 de 07 de 2017). Planea: Tensiones y Contradicciones. *La Jornada*.
- Capdevilla, B. I. (6 de 2013). Serious Game Architecture and Design: modular component-based entity system framework to support systemic modeling and design in agile serious game developments. *Tesis Doctoral*. Paris, Francia.
- Fullerton, T. (2008). *Game Design Workshop: A Playcentric Approach to Creating Innovative Games, Second Edition*. Morgan Kauffman Publishers.
- GameLearn. (6 de 3 de 2017). *GameLearn*. Obtenido de GAMELEARN, S.L.: <https://www.game-learn.com/lo-que-necesitas-saber-serious-games-game-based-learning-ejemplos/>
- INEE. (2018). *PLANEA - Resultados Nacionales*. México: Instituto Nacional para la Evaluación de la Educación.
- ITM. (2015). Reporte Edutrends. *Observatorio de Inovacion Educativa del Tecnologico de Monterrey*.
- Johnson, J. (2015). Changing the Learning Environment in the College of Engineering and Applied Science Using Challenge Based Learning. *International Journal of Enginnering Pedagogy*.
- Marcano, B. (2008). Juegos serios y entrenamiento en la sociedad digital. *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*.
- Medina, A. C. (2007). *Nuevas tecnologías para la educación en la era digital*. Piramide.
- Piaget, J. (1978). *La representación del mundo en el niño*. Madrid: Morata.

Ricardo Emmanuel Gutiérrez-Hernandez, F. J.-A. (2013). Arquitectura de Software para Juegos Serios con Aspectos Culturales: Caso de estudio en un Videojuego para Formulas Temperatura.

Santibáñez, L. (2015). ¿Qué nos dice Planea? *Nexus*.

UNESCO. (1984). Glosario de Terminos de Tecnologia Educativa. *Perspectivas: revista trimestral de educación comparada*.

Urquidí, M., Cristina, A., & Aznar, C. T. (2015). Juegos serios como instrumento facilitador del aprendizaje: evidencia empirica. *Opción*.

Walid Mestadi, K. N. (2018). An Assessment of Serious Games Technology: Toward an Architecture for Serious Games Design. *International Journal of Computer Games Technology*.

## Notas biográficas



**Ing. Carlos Enrique Flores Méndez** recibió el título de Ingeniero en Sistemas y Comunicaciones por la Universidad Autónoma del Estado de México. Actualmente, se encuentra estudiando la Maestría en Ciencias de la Computación en el Centro Universitario UAEM Valle de México de dicha institución. Sus áreas de interés se encuentran en el desarrollo de juegos serios, inteligencia artificial y desarrollo de software educativo.



**Dra. Maricela Quintana López** es Profesora de Tiempo Completo del Centro Universitario UAEM Valle de México de la Universidad Autónoma del Estado de México. La Dra. Quintana es Ingeniera en Sistemas Computacionales, y tiene la Maestría y el Doctorado en Ciencias Computacionales con especialidad en Inteligencia Artificial por el Tecnológico de Monterrey. Cuenta con el reconocimiento de Perfil Deseable PRODEP de la SEP en México. Ha dirigido tesis de maestría y licenciatura, ha sido conferencista magistral y ponente en diversos eventos. Ha publicado artículos en las áreas de Inteligencia Artificial y Minería de Datos relacionados con sus proyectos de Investigación.



**Dr. Héctor Rafael Orozco Aguirre** es Profesor de Tiempo Completo del Centro Universitario UAEM Valle de México. Obtuvo su Maestría en 2006 y Doctorado en 2010 en el Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Guadalajara. Como parte de sus estancias de investigación en el extranjero, de 2007 a 2008 estuvo en el Laboratorio de Realidad Virtual de la Escuela Politécnica Federal de Lausana en Suiza, y de 2011 a 2012 en el Instituto de Innovación en Medios de Comunicación de la Universidad Tecnológica de Nanyang en Singapur. En 2011, fue galardonado por la Sociedad Mexicana de Inteligencia Artificial a la Tercera Mejor Tesis de Doctorado a nivel nacional. Actualmente, trabaja en Proyectos de Investigación referentes a Simulación de Estrategias de Predicción y Anticipación de Crimen para su Control y Reducción, Tutores Virtuales para la Mejora de la Enseñanza Educativa, así como Análisis, Modelado y Simulación de Tráfico Vehicular y Comportamiento Peatonal. Ha dirigido tesis de maestría y licenciatura, Ha sido conferencista magistral y ponente en diversos eventos y ha publicado artículos y capítulos de libro en las áreas de Inteligencia Artificial Aplicada e Inteligencia Computacional. Cuenta con el reconocimiento de Perfil Deseable PRODEP de la SEP en México.



**Dra. Ivonne Rodríguez Pérez** Licenciada en Matemáticas Aplicadas y Computación egresada de la Universidad Nacional Autónoma de México, tiene la Maestría en Ciencias de la Educación en la Universidad del Valle de México, Doctorado en Educación en el Centro de Estudios Superiores en Educación, es Profesora de Tiempo Completo del Centro Universitario UAEM Valle de México de la Universidad Autónoma del Estado de México. Cuenta con el reconocimiento de Perfil Deseable PRODEP de la SEP en México.



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 2.5 México.



*Recibido 08/06/2019*

*ReCIBE, Año 8 No. 2, Noviembre 2019*

*Aceptado 02/09/2019*

## **SDR y GNU Radio como plataforma para un laboratorio de comunicaciones digitales**

**Rogelio Jaimes Rico<sup>1</sup>**  
roger-jr97@hotmail.com

**Saul Lazcano Salas<sup>1</sup>**  
slazcanos@uaemex.mx

<sup>1</sup>Centro Universitario UAEM Valle de México  
Universidad Autónoma del Estado de México

**Resumen:** SDR (Software Defined Radio) o radio definido por software es una tecnología flexible compuesta por una parte hardware (radio frecuencia) y otra parte basada en software (procesado en banda base). El presente trabajo tiene como fin el análisis de SDR como base para la implementación de un laboratorio de comunicaciones digitales en licenciaturas vinculadas a las telecomunicaciones, tales como Ingeniería en Computación o Ingeniería en Sistemas y Comunicaciones. Para ello, se analizan algunos temas básicos como muestreo, filtrado, modulación y la factibilidad de implementación de los mismos en un esquema SDR, mediante un radio modelo HackRf One en un entorno de GNU Radio.

**Palabras clave:** GNU Radio, comunicaciones digitales, SDR.

**Abstract:** SDR (Software Defined Radio) or software-defined radio is a flexible technology composed of a hardware part (radio frequency transceiver) and other software-based part (band pass processing). The purpose of this paper is to analyze the use of SDR as the basis for the implementation of a digital communications laboratory in degrees related to telecommunications, such as Computer Engineering or Systems and Communications Engineering. In this way, some basic topics such as sampling, filtering, modulation and the feasibility of implementing them in an SDR scheme are analyzed, using a HackRF One Radio under a GNU Radio environment.

**Keywords:** GNU Radio, digital communications, Software Defined Radio.

## 1. Introducción

GNU Radio (The GNU Radio Foundation, 2019), es una herramienta de desarrollo con licencia tipo GLP (General Public License), que permite, mediante bloques de procesamiento de señales, la construcción y simulación de sistemas de comunicaciones, lo anterior puede ser a nivel simulación o a nivel implementación con la ayuda de radios programables. La tarea de los radios es trabajar a nivel radio frecuencia, mientras el procesamiento en banda base de la señal se realiza mediante el software.

Esta combinación de elementos hardware y software permiten la creación de diferentes esquemas de comunicaciones a partir de un único dispositivo hardware (radio), abriendo la puerta para que en la actualidad, los usuarios de este tipo de tecnología sea muy extensa, desde radio aficionados hasta para realizar investigaciones de frontera en diversas áreas vinculadas a las comunicaciones inalámbricas en general, tales como codificación de canal (Giard, Sarkis, Leroux, & et.al., 2018), redes programables (Macedo, Guedes, Vieira, Vieira, & Nogueira, 2015; Foukas, Nikaein, Kassem, Mahesh, & Kontovasilis, 2016), antenas (Abari, 2016), educación, por citar algunas áreas (Wyglinski, Orofino, Ettus, & Rondeau, 2016).

Por otro lado, existen diversos radios SDR con diferentes características, entre los que destaca, por su relación calidad / precio, el dispositivo HackRF One, que junto al software GNU Radio permite realizar conversiones A/D (Analógico - Digital) y D/A (Digital -Analógico), con una tasa de muestreo de hasta 20 millones de muestras por segundo, trabajar con señales en el rango de 1 MHz – 6 GHz además de capturar, analizar y transmitir frecuencias de radio (Leahy Center for Digital Investigation, 2017).

En conjunto, software (entorno GNU Radio) y hardware (equipo HackRF One), constituyen una herramienta sumamente atractiva, particularmente para fines de docencia dentro del área de sistemas de comunicaciones digitales, ya que permite la construcción y el modelado de elementos tales como conversores analógico-digital y digital analógico, filtros diversos, receptores y emisores completos, entre otros, facilitando enormemente la comprensión de dichos temas por parte de los alumnos.

## 2. Conceptos básicos

ADC (Convertidor Digital Analógico): transforma una señal eléctrica analógica en un número digital equivalente.

DAC (Convertidor Analógico Digital): transforma un número digital en una señal eléctrica analógica (Huircan, 2003).

Teorema de Nyquist: reconstruye una señal analógica en base a la conversión analógica digital, para que esto se cumpla la señal analógica debe ser limitada en banda y la frecuencia de muestreo debe ser igual o mayor al doble del ancho de banda de la señal, esto si se da por hecho que la señal analógica esta en banda base, este teorema se puede ver reflejado en la conversión analógica-digital y se encuentra representado en la siguiente expresión:

$$F_s \geq 2F_c \quad (1)$$

Donde:

F<sub>s</sub>: frecuencia de muestreo (Sampling).

F<sub>c</sub>: frecuencia de la información – ancho de banda (Carrier).

GNU Radio: se define como un entorno de desarrollo libre y abierto (The GNU Radio Foundation, 2019); los bloques de procesamiento de señales son de código abierto (C++ o Python). Se enfoca de manera directa en ser una herramienta de apoyo para el trabajo con radios SDR.

Su manejo es simple e intuitivo, esto hace que sea sumamente amigable para todo tipo de usuario, característica particularmente útil para nuevos usuarios. Adicionalmente, cuenta con una comunidad de usuarios que brinda soporte y consejos en foros especializados (FAQ-GNU Radio, 2016).

La interfaz gráfica puede ser dividida en 4 partes, como se muestra en la figura 1:



**Figura 1.** Interfaz gráfica en GNU Radio con todos sus componentes.

1) Barra de herramientas: espacio configurable para las funciones esenciales, como ejecutar el sistema, guardar proyecto, nuevo espacio de trabajo, copiar, pegar, retroceder (estado previo del proyecto), buscar, habilitar y deshabilitar bloques, por mencionar las más importantes y usadas.

2) Espacio de trabajo: espacio en blanco donde se pueden diseñar sistemas conformados por bloques conectados unos de otros, por defecto, al abrir un nuevo proyecto estarán dos bloques (top\_block y samp\_rate).

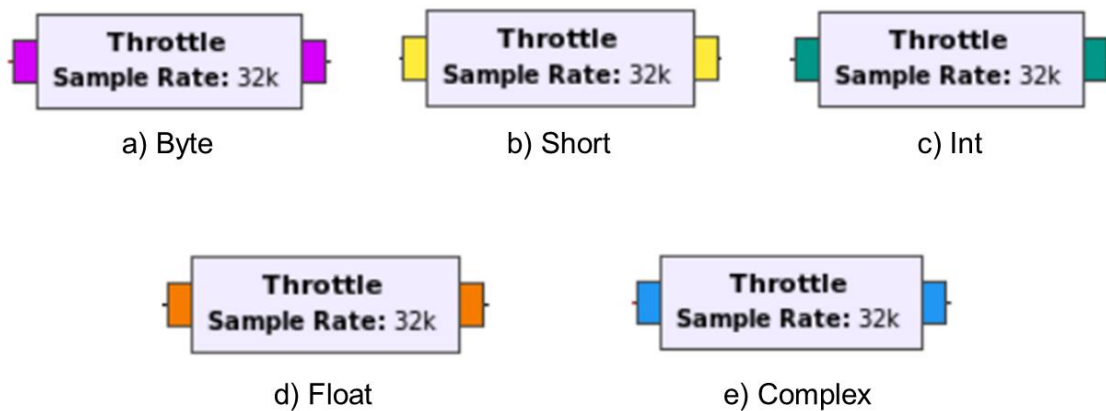
3) Biblioteca de bloques: se encuentran todos los bloques de procesamiento con los que el software está previamente conformado, cada bloque se encuentra ordenado acorde a la función que realiza, basta con acudir a un apartado y desplegar la lista de bloques que contiene.

4) Consola: ubicado en la parte inferior de la interfaz gráfica, en esta se muestra el estado de la ejecución del sistema.

HackRF One: el dispositivo de tipo SDR (Software Defined Radio) HackRF One (Great Scott Gadgets, 2016), tiene la capacidad de digitalizar señales de radio que son recibidas o transmitidas por otros dispositivos, trabaja en un rango de frecuencia desde 1MHz hasta 6GHz, lo cual permite abarcar dispositivos que operan con Bluetooth, radio FM, comunicación de campo cercano (del término en inglés NFC, Near Field Communication) y tecnologías de telefonía móvil.

### 3. Diseño de sistemas en GNU Radio

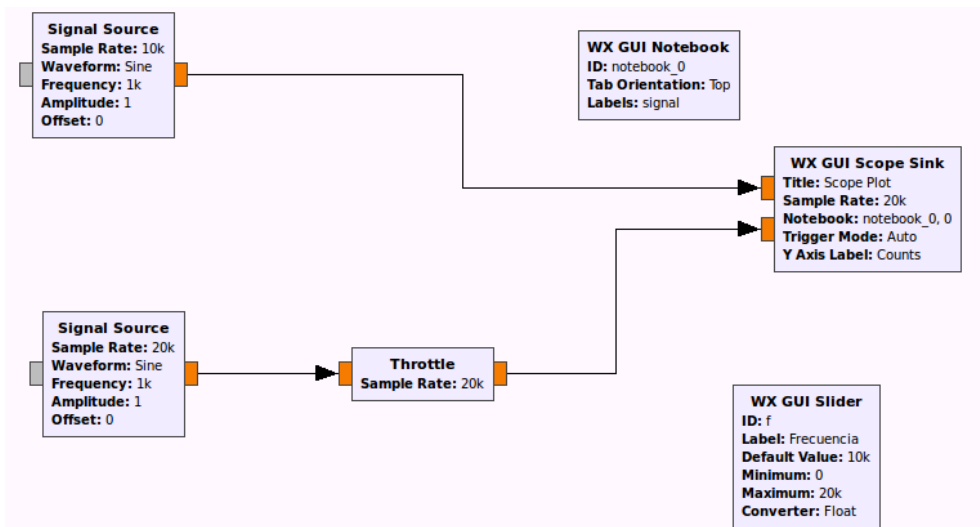
Como se ha mencionado, GNU Radio cuenta con una interfaz intuitiva para la creación de sistemas de radiofrecuencias, cada bloque se caracteriza por el tipo de datos de entrada y datos de salida que soporta. Los tipos de datos básicos que GNU Radio soporta son: byte, short, in, floating point y complex, los cuales se distinguen a partir de colores: morado, amarillo, verde, naranja y azul, como se muestra en la figura 2.



**Figura 2.** Tipos de datos y sus colores de identificación. Fuente: (FAQ-GNU Radio, 2016).

A continuación se muestran algunos ejemplos de sistemas realizados en GNU Radio en conjunto con el radio HackRF One, cada uno con una función específica:

**Convertor analógico-digital:** convierte una señal analógica en una señal digital, y es capaz de adaptarse a los dispositivos de procesos involucrados en un sistema de transmisión, se muestra en la figura 3.



**Figura 3.** Esquema general de un convertor analógico – digital.

**Receptor FM:** este diagrama en uso conjunto al dispositivo SDR HackRF One tiene la capacidad de sintonizar estaciones de radio FM y al mismo tiempo graficar la transformada de Fourier de dicha señal a manera de analizador de espectros. Este diagrama se muestra en la figura 4.

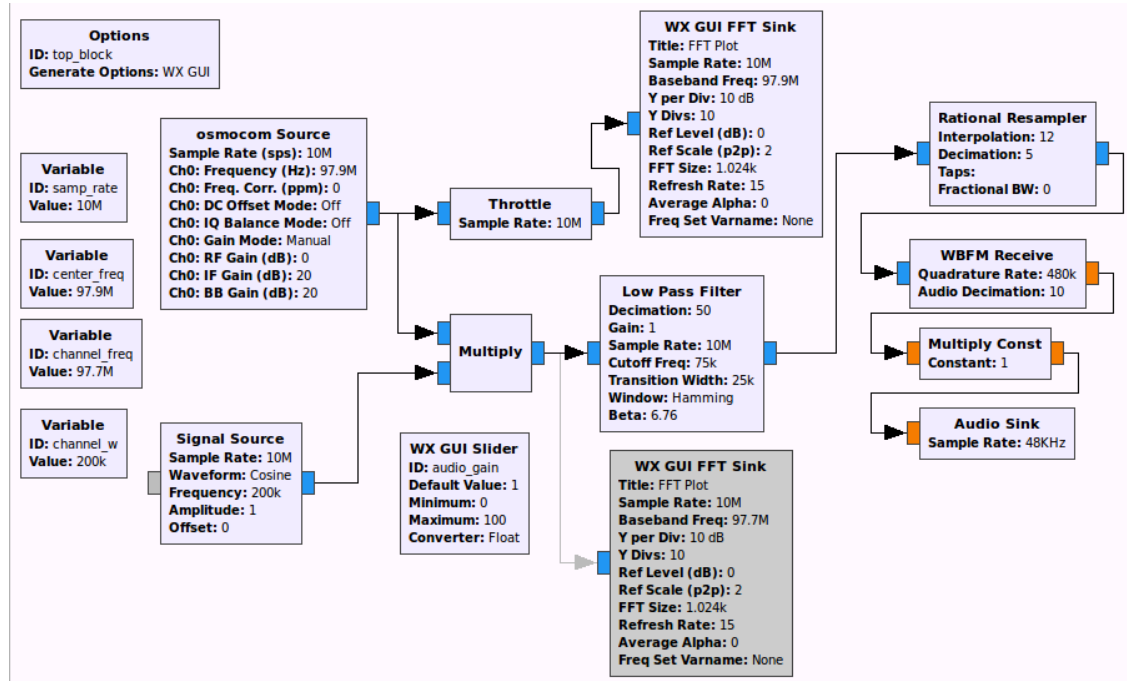


Figura 4. Sistema receptor de señales FM.

**Receptor FM de 2 frecuencias:** variante del sistema anterior, en donde se añade la recepción de una segunda señal de FM y al igual que el caso anterior, se grafica la transformada de Fourier de la señal a manera de analizador de espectros. El esquema de bloques se muestra en la figura 5.

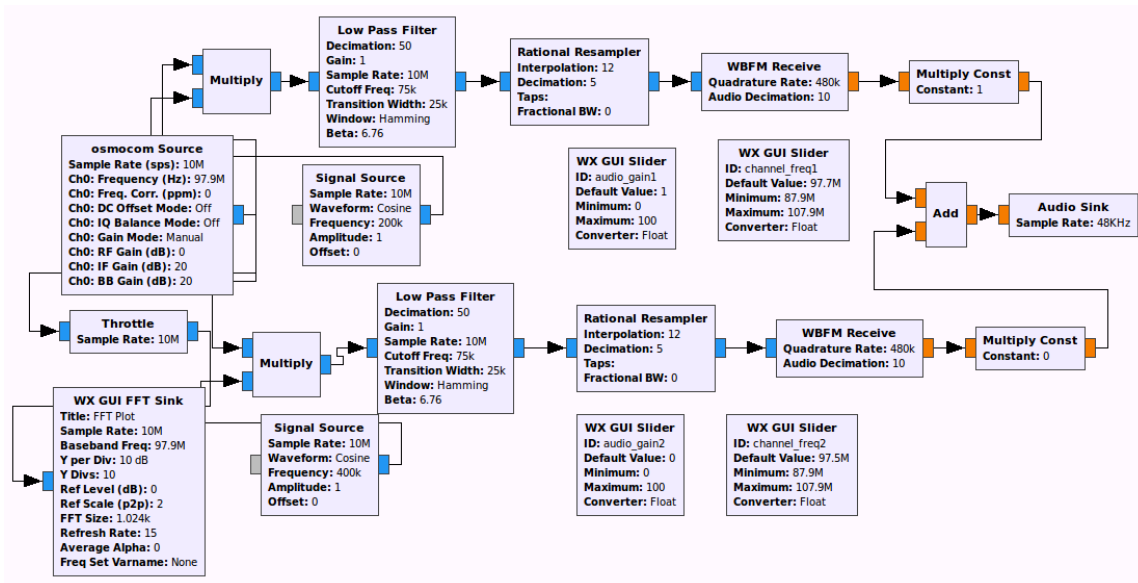
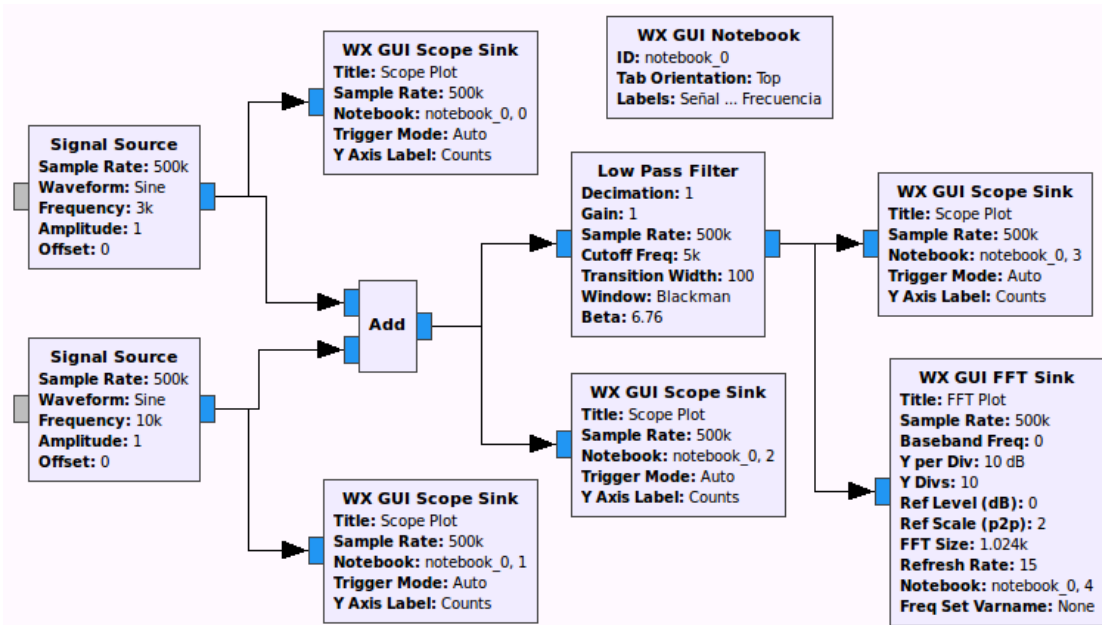


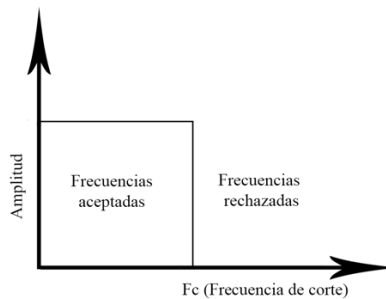
Figura 5. Sistema receptor de dos señales FM.

**Filtro pasa bajo:** tiene la finalidad de permitir el paso a señales con una frecuencia por debajo de un rango estipulado, por otro lado, rechaza aquellas que superan ese rango. El esquema se muestra en la figura 6.



**Figura 6.** Diseño de un filtro pasa bajo.

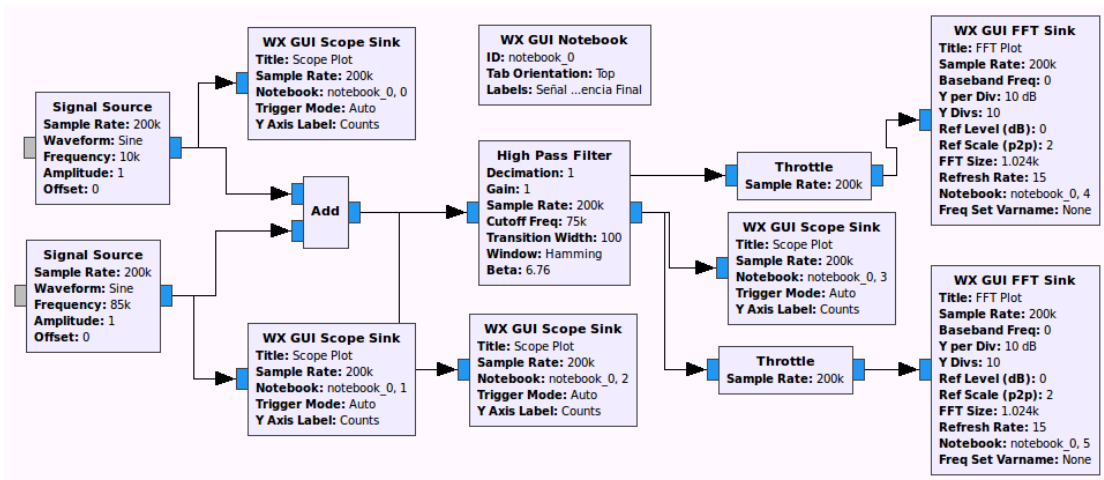
La figura 7 define el comportamiento ideal de un filtro paso baja, el cual permite el paso de señales por debajo de la denominada frecuencia de corte y elimina las señales por arriba de dicha frecuencia.



**Figura 7.** Representación del funcionamiento de un filtro pasa bajo ideal.

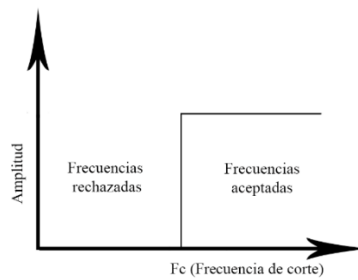
**Filtro paso alto:** de manera opuesta al comportamiento de un filtro paso baja, el filtro paso alto permite el paso de señales cuya frecuencia sea superior a una determinada frecuencia umbral predeterminada; aquellas señales con frecuencia inferior a la frecuencia umbral son atenuadas. El esquema general de la construcción de un filtro paso altas se muestra en la figura 8.





**Figura 8.** Diseño de un filtro pasa alto.

Como se puede apreciar en la figura 9, la forma en que actúa un filtro pasa alto ideal, dejando pasar solo aquellas frecuencias que están por encima de la frecuencia de corte.



**Figura 9.** Representación del funcionamiento de un filtro pasa alto ideal.

**Filtro pasa banda:** Permite el paso únicamente a aquellas señales que se encuentren en un determinado rango, limitado por una frecuencia umbral inferior y una frecuencia de corte superior (el valor más alto que puede aceptar), el diagrama de bloques de la construcción de un filtro paso banda en GNU Radio se muestra en la figura 10.

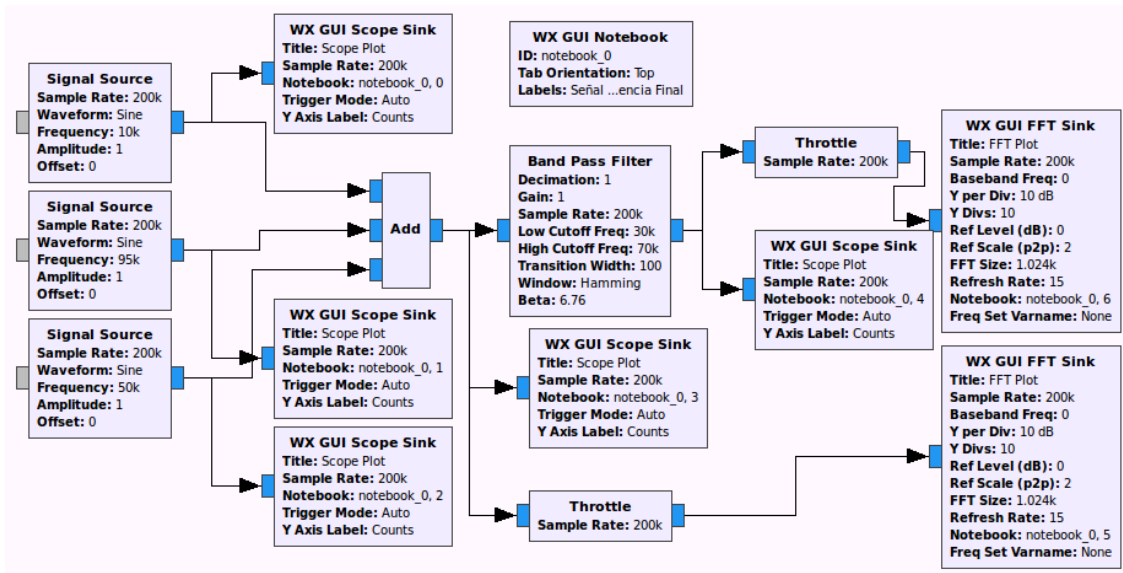


Figura 10. Diseño de un filtro pasa banda.

La figura 11 presenta el comportamiento de un filtro pasa banda ideal, dejando pasar solo aquellas frecuencias que se encuentran entre  $F_1$  y  $F_2$ , rechazando todas aquellas fuera de estos rangos.

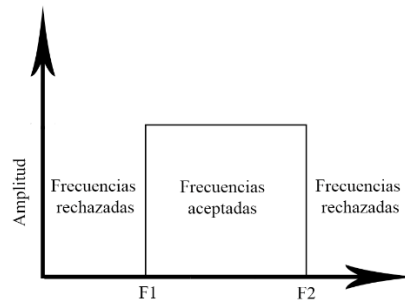
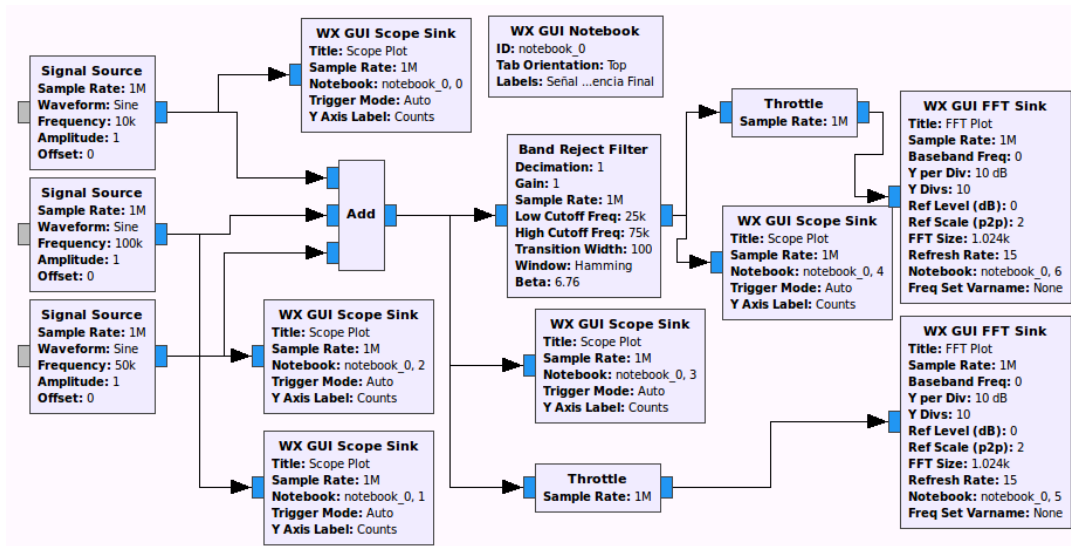


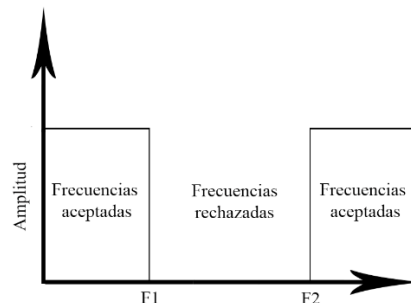
Figura 11. Representación del funcionamiento de un filtro pasa banda.

**Filtro rechazada banda:** de manera análoga al filtro paso banda, este filtro requiere de igual manera una frecuencia inferior y una superior, sin embargo, este filtro atenúa todas las señales comprendidas en el rango antes definido. El diagrama de bloques en GNU Radio se muestra en la figura 12.



**Figura 12.** Diseño de un filtro rechaza banda.

Como se puede apreciar en la figura 13, la forma en que actúa un filtro rechaza banda, dejando pasar solo aquellas frecuencias por debajo de  $F_1$  y encima de  $F_2$ , rechazando todas aquellas fuera de estos rangos.



**Figura 13.** Representación del funcionamiento de un filtro rechaza banda ideal.

#### 4. Conclusiones y trabajo futuro

SDR es una tecnología que presenta un área de oportunidades muy extenso tanto en docencia, aplicaciones a nivel amateur y experimental y por supuesto, en la investigación.

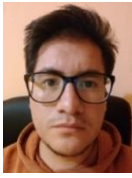
La combinación de radios con una buena relación calidad / precio con un entorno de desarrollo abierto y con una gran comunidad de soporte como lo es GNU Radio, da como resultado una excelente combinación que genere una gran inercia en beneficio de los alumnos, facilitando la comprensión de temas asociados a comunicaciones digitales. En este sentido, un laboratorio completo de comunicaciones digitales basado en tecnología SDR, abriría las puertas no sólo para cubrir lo relacionado a la docencia, sino permitiría generar una masa crítica en los alumnos para ir más allá y apostar por la creación de grupos de trabajo estudiantiles con un perfil muy sólido en comunicaciones inalámbricas en general.

Como trabajo a ser desarrollado, es el estudio de esquemas de codificación de canal enfocados a canales inalámbricos, particularmente códigos polares y su decodificación mediante algoritmos basados en redes neuronales.

## REFERENCIAS

- Abari, O. (2016). A millimeter wave software defined radio platform with phased arrays. *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking (MobiCom '16)* (págs. 419-420). New York: ACM. doi:10.1145/2973750.2985258
- FAQ-GNU Radio. (1 de Junio de 2016). *FAQ-GNU Radio*. Recuperado el 10 Agosto de 2019, de FAQ-GNU Radio: <https://wiki.gnuradio.org/index.php/FAQ>
- Foukas, X., Nikaiein, N., Kassem, M., Mahesh, K., & Kontovasilis, K. (2016). FlexRAN: A Flexible and Programmable Platform for Software-Defined Radio Access Networks. *Proceedings of the 12th International on Conference on emerging Networking EXperiments and Technologies (CoNEXT '16)* (págs. 427-441). New York: ACM. doi:10.1145/2999572.2999599
- Giard, P., Sarkis, G., Leroux, C., & et.al. (Mayo de 2018). Low Latency Software Polar Decoders. *Journal of Signal Processing Systems*, 761-775. doi:10.1007/s11265-016-1157-y
- Great Scott Gadgets. (01 de Enero de 2016). *HackRF - Great Scott Gadgets*. Recuperado el 10 Agosto de 2019, de Great Scott Gadgets: <https://greatscottgadgets.com/hackrf/>
- Huircan, J. (1 de Enero de 2003). *Web Quidel Universidad de la Frontera*. Recuperado el Agosto de 2019, de Web Quidel: [http://quidel.inele.ufro.cl/~jhuircan/PDF\\_CTOSII/ad03.pdf](http://quidel.inele.ufro.cl/~jhuircan/PDF_CTOSII/ad03.pdf)
- Leahy Center for Digital Investigation. (1 de Enero de 2017). *Champlain College*. Recuperado el 10 Agosto de 2019, de Leahy Center for Digital Investigation: [https://www.champlain.edu/Documents/LCDI/HackRF%20One%20Tutorial\\_F2017%20-%20Report.docx.pdf](https://www.champlain.edu/Documents/LCDI/HackRF%20One%20Tutorial_F2017%20-%20Report.docx.pdf)
- Macedo, D. F., Guedes, D., Vieira, L., Vieira, M., & Nogueira, M. (2015). Programmable Networks, From Software-Defined Radio to software-Defined Networking. *IEEE Communications Surveys & Tutorials*, 17(2), 1102-1125. doi:10.1109/COMST.2015.2402617
- The GNU Radio Foundation. (1 de 1 de 2019). *GNU Radio - The Free & Open Source Radio Ecosystem*. Recuperado el 10 Agosto de 2019, de GNU Radio: <https://www.gnuradio.org/>
- Wygłinski, A. M., Orofino, D. P., Ettus, M. N., & Rondeau, T. W. (Enero de 2016). Revolutionizing software defined radio: case studies in hardware, software, and education. *IEEE Communications magazine*, 54(1), 68-75. doi:10.1109/MCOM.2016.7378428

## Notas biográficas



**Rogelio Jaimes Rico** estudiante de noveno semestre de la carrera de Ingeniería en Sistemas y Comunicaciones del Centro Universitario UAEM Valle de México de la Universidad Autónoma del Estado de México, México. Su línea de acentuación es telecomunicaciones y está involucrado en trabajos de investigación relacionados a SDR (Software Defined Radio).



**Saul Lazcano Salas** es Profesor de Tiempo Completo del Centro Universitario UAEM Valle de México de la Universidad Autónoma del Estado de México, México. Egresó de la licenciatura en Ingeniería en Telecomunicaciones por la Universidad Nacional Autónoma de México, México, hizo su Máster en Gestión de las Telecomunicaciones por la Escuela de Organización Industrial en España y su Doctorado en Ingeniería, área de Telecomunicaciones por la Universidad Nacional Autónoma de México, México. Actualmente, desarrolla investigación en el área de codificación de canal y áreas afines. Ha dirigido tesis de maestría y licenciatura, ha sido conferencista ponente en diversos eventos y ha publicado artículos y capítulos de libro en las áreas de Inteligencia Artificial Aplicada e Inteligencia Computacional. Cuenta con el reconocimiento de Perfil Deseable PRODEP de la SEP en México.



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 2.5 México.

# Análisis de sentimientos en videojuegos

Daniel Iván Angeles Gómez<sup>1</sup>  
angodi228@gmail.com

Maricela Quintana López<sup>1</sup>  
mqintanal@uaemex.mx

<sup>1</sup>Centro Universitario UAEM Valle de México  
Universidad Autónoma del Estado de México

## Resumen.

Los videojuegos multijugador han existido desde la década de los setenta, pero no es hasta mediados de los noventa, con la evolución de las redes que permiten la comunicación desde cualquier parte del mundo, que adquieren la característica de jugarse en línea, dándole otra dimensión humana al juego ya que se requiere una manera de comunicarse para formular estrategias o simplemente para platicar mientras se juega. La comunicación se puede realizar vía voz o texto a través de una interfaz, y a lo largo del juego se pueden tener conversaciones que van desde las casuales, que tienen que ver con aspectos generales del juego, hasta las agresivas, que incluyen insultos a los demás jugadores. En este artículo se muestran los resultados de aplicar el análisis de sentimientos a 75 conversaciones extraídas de partidas de los videojuegos League of Legends, Overwatch y Heroes of the Storm, 25 están en idioma español, y 50 en inglés; las conversaciones son de longitud variable, la más extensa contiene 440 palabras, mientras que la más pequeña contiene solo 50 palabras. Utilizando Semantría para Excel, un API del software Lexalytics, se creó un diccionario que pudiera ser utilizado para la clasificación de las conversaciones. Las clases se definieron como positivas, negativas o neutrales. Comparando el resultado con la clase dada por el experto, 47 de 75 conversaciones fueron clasificadas correctamente, dando un 62.6% de certeza a la clasificación, en búsqueda de aumentar la certeza de la clasificación se hizo uso de el algoritmo C4.5, en el cual los resultados mostraron el entorno de los implicados en las conversaciones, la certeza se considera podría mejorarse al incluir más conversaciones y un mejor balance en las clases.

**Palabras clave:** Análisis de sentimientos, clasificación, videojuegos.

## Abstract

Multiplayer videogames exist since the 70's (seventies), but it isn't until mid 90's (nineties) with the evolution of networks that allow communication from anywhere in the world, that videogames become online games, giving them a new human dimension by requiring a way to communicate in order to create strategies, or just talk while playing. The communication can be via voice or text lines within an interface, and throughout the game, conversations can go from casual topics related to the general aspects of the game, to aggressive conversations insulting other players. This article shows the results of a Sentiments Analysis of 75 conversations extracted from the videogames: League of Legends, Overwatch and Heroes of the Storm; 25 in Spanish language, and 50 in English. The conversations are of variable length, the longest has 440 words, while the shortest only has 50 words. Using Semantría for Excel, an API software of Lexalytics, a lexicon dictionary was created in order to classify the conversations.

The classes were defined as positive, negative or neutral. Comparing these results with the given ones by an expert, 47 of the 75 conversations were correctly classified, getting a 62.6% classification accuracy, which can be improved using the C4.5 algorithm, the results showed the environment in the conversations players, the accuracy can be better with more conversations and a higher class balance.

**Keywords:** Sentiment Analysis, Classification, Videogames.

## 1 Introducción

Los videojuegos han tenido un creciente impacto en la industria del entretenimiento gracias a los dispositivos móviles, ya que estos permiten el uso de diferentes tipos de juegos, que van desde rompecabezas (“puzzles”), sin un fin específico (“arcade”), de estrategia (“RTC” Real Time Combat), hasta juegos donde su historia es inmersiva llegando a los niveles de un filme cinematográfico. Su gran diversidad ha permitido que personas de todas las edades los jueguen (Harris, 2014).

Hoy en día, los videojuegos, han llegado a formar parte fundamental en la rutina diaria de muchas personas, ya que están presentes como medios de entretenimiento y en algunos casos como vida laboral, esto incluye, a los que se dedican a desarrollar los videojuegos y a los jugadores profesionales que participan en diferentes eventos obteniendo una remuneración (Steven, 2016).

Si bien desde un principio, han existido juegos para uno o varios jugadores, no es hasta que la tecnología, particularmente las redes, evolucionaron al punto en el que la comunicación podía realizarse desde cualquier parte del mundo, que los juegos multijugador adquieren la característica de poderse jugar en línea (“online”), dándole otra dimensión humana al juego ya que se requiere una manera de comunicarse para formular estrategias o simplemente para platicar mientras se juega. La comunicación se puede realizar vía voz o texto a través de una interfaz, y a lo largo del juego se pueden tener conversaciones que van desde las casuales, que tienen que ver con aspectos generales del juego, hasta las agresivas, que incluyen insultos a los demás jugadores.

Dada la gran variedad de gente que interactúa en los videojuegos, existe una inmensa variedad de conversaciones que reflejan cómo se comportan los jugadores a lo largo de una partida en diferentes circunstancias del juego, desde cuándo van ganando hasta cuando la derrota es inminente, por lo que dichas conversaciones pueden analizarse utilizando la minería de opinión o análisis de sentimientos.



A diferencia de otras aplicaciones en donde todas las conversaciones tienden a ser de dominio público, en el caso particular de los videojuegos se carece del acceso directo a las conversaciones, por ende, es más complicado el extraer la información. En algunas ocasiones, los juegos varían en algunos aspectos dependiendo de la región, entre estos, el doblaje del juego, la ambientación y posibles temas sensibles, por ejemplo, el juego Wolfenstein tiene referencias al régimen nazi, por lo que en Alemania los símbolos o personajes contienen variaciones. Otro aspecto relacionado con la región es el servidor donde se alojan las partidas.

En este trabajo, se presentan los resultados del análisis de sentimientos realizado utilizando las conversaciones de tres juegos que tienen una versión similar para todo el mundo y solo varían en el doblaje: League of Legends (“LOL”), Overwatch (“OVW”) y Heroes of the Storm (“HOTS”), en el caso del primero se ocupará el servidor LAN (Latino América) y en los últimos dos el servidor de las “AMERICAS” que ocupa todo el continente americano.

El documento está organizado como sigue, en la sección 2 se presenta el análisis de sentimientos y la metodología que se utiliza. En la sección 3 se presentan las etapas realizadas para el análisis de sentimientos en videojuegos. Finalmente en la sección 4, se presentan las conclusiones y trabajo futuro.

## **2 Análisis de sentimientos**

El análisis de sentimientos, SA, por sus siglas en inglés Sentiment Analysis, es el estudio de conversaciones donde se incluyen opiniones, actitudes y emociones realizado por computadora, puede ir enfocado a conversaciones individuales, hacia un tema o ciertos tipos de eventos, el SA se desempeña en diferentes niveles, pueden ser, a nivel de documento, oración o aspecto, en el primer nivel refleja los sentimientos expresados en un documento que son clasificados como positivos o negativos, en el segundo caso se realiza en base a oraciones que llegan a ser de poca longitud, existe mucha similitud en ambos casos debido a que las oraciones son subconjuntos de un documento, por último se tiene el análisis por aspecto, en los aspectos se intenta encontrar palabras claves que puedan ser de gran impacto en una oración o documento (Medhat, Hassan, & Korashy, 2014).

Los campos más explotados dentro del análisis de sentimientos están comúnmente relacionados con las redes sociales y con las reseñas de productos o servicios, los estudios presentan la opinión de las personas en base a un tema o dirigido a un producto, comúnmente clasificado en Positivo, Neutral o Negativo, los análisis que han tenido mayor mención se han desarrollado alrededor de Twitter, cada estudio tiene sus propias características como puede ser la metodología aplicada, las herramientas o la cantidad de datos.

Dentro del campo del análisis de sentimientos, se encuentra el SemEval-2016 Task 4: Sentiment Analysis in Twitter, un estudio que pretende averiguar los temas positivos o negativos que han surgido en Twitter. Una de las principales costumbres en este tipo de análisis es tener dos parámetros para los sentimientos: NEGATIVO y POSITIVO, sin embargo, SemEval-2016 Task 4, debido a la gran cantidad de datos e información que manejan, añade tres evaluaciones más, teniendo como parámetros finales: MUY POSITIVO, POSITIVO, NEUTRAL, NEGATIVO y MUY NEGATIVO, esto con el fin de que al tener más evaluaciones se tiene un mayor grado de exactitud al momento de clasificar las conversaciones, se utilizaron diferentes herramientas para comparar la certeza de cada una. El peor resultado que se registró fue de 37.4% de asertividad mientras que el más alto obtuvo una puntuación de 64.6% (Navok, Ritter, Rosenthal, Sebastiani, & Stoyanov, 2016).

Un año más tarde se realizó nuevamente el estudio SemEval-2017 Task 4, se ocuparon metodologías y procedimientos similares. Las principales modificaciones que se realizaron fueron las actualizaciones de algunas herramientas, se añadieron nuevos sistemas y hubo incorporación de diferentes lenguajes como el arábico o el español, dentro de la clasificación se observaron diferentes mejorías en la clasificación, donde la más baja obtuvo 20.6% de asertividad mientras que el más alto consiguió un total de 89.7% (Rosenthal, Farra, & Nakov, 2017).

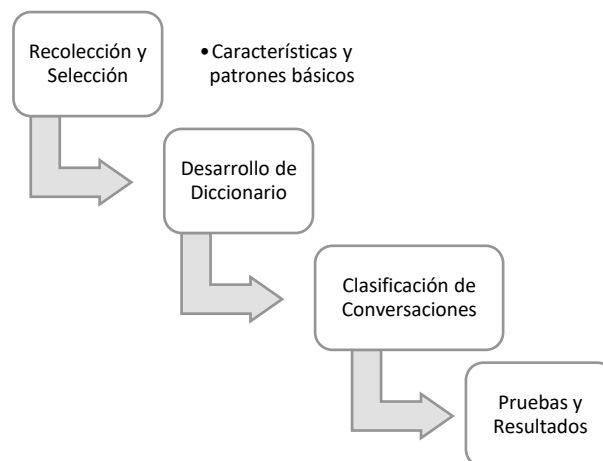
Existen muchas formas de evaluar una conversación o frases dentro del SA, la más usual es evaluar toda la conversación sin modificar en ningún momento su estructura, sin embargo, también se pueden retirar cierta parte de la información como pre-procesamiento, tal como demostraron (Mat Zin, Mustapha, Azmi Murad, & Mohd Sharef, 2017) en su artículo donde analizan reseñas de películas. Parte de la metodología, fue retirar contenido para centrarse en las palabras o frases en las conversaciones, el contenido retirado fueron artículos, palabras sin sentido, fechas, caracteres especiales, números y palabras con menos de tres caracteres, de esta manera lograron aumentar la certeza del análisis.

En el campo de los videojuegos, es raro encontrar un estudio de esta naturaleza, normalmente van enfocados al aspecto psicológico o al desarrollo. Lo que se ha realizado son estudios en base a las reseñas de los juegos, que se hacen tanto de manera profesional como de manera casual (las escritas por los jugadores). En el artículo de (Strååt & Verhagen, 2017) utilizan el análisis de sentimientos para clasificar las reseñas en malas, neutrales o buenas con base a 3 características de los juegos, las cuales son la historia, el combate y los personajes. Dentro de los resultados encontraron que la mayoría de las críticas se enfocaron en la historia, los resultados muestran que la mayoría de las reseñas están enfocadas en la historia y están reseñadas como malas historias.

## 2.1 Metodología

Dado el amplio campo de aplicación del análisis de sentimientos, se realizan estudios en base a las redes sociales y a las opiniones de los consumidores, (Fang & Zhan, 2015) realizó un estudio utilizando las opiniones de los clientes en la plataforma de Amazon, se utiliza una metodología de 3 fases (Recolección, Identificación de sentimientos y resultados), y al terminar se aplican algunos algoritmos para vislumbrar los principales comportamientos de las reseñas que los clientes han publicado.

Debido a que el SA es un proceso de clasificación, se deben seguir una serie de pasos para realizarlo con éxito, algunos autores manejan metodologías diferentes pero con bases similares, en el caso de (Medhat, Hassan, & Korashy, 2014) mencionan 5 pasos que incluyen la revisión del material, la identificación de los sentimientos, la selección de características, la clasificación de sentimientos y los resultados. En este trabajo, se han realizado adecuaciones a esta metodología para enfocarla a las características del caso de estudio, las cuales se muestran en la figura 1.



**Figura 1.** Metodología para el SA basado en el modelo que presenta (Medhat, Hassan, & Korashy, 2014)

La primera etapa consiste en la extracción de la información, seleccionando aquellas muestras útiles para el estudio, seguido del uso de un software para desarrollar el diccionario de palabras en el que se añade un peso informacional a los datos con base al tema o evento. Una vez que el software realiza la evaluación de la información, esta se clasifica en negativo, positivo o neutral y con los resultados se evalúan los aciertos obtenidos de la misma y con ello se procede a experimentar con otros valores para mejorar su precisión.

## **2.2 Herramientas**

Existen diferentes herramientas para el análisis de sentimientos, en muchas ocasiones se presentan como herramientas enfocadas en el aspecto mercantil, es decir, en base a una marca se evalúan los sentimientos u opinión de la misma en determinadas redes sociales y no permiten la modificación de ciertos criterios, como el umbral de decisión o el diccionario que utilizan, un ejemplo de este tipo de sistemas es Brandwatch, una aplicación que permite el análisis de sentimientos de una o varias marcas, comparando las opiniones en redes sociales y así saber que marca es del agrado del público en general, usualmente estas redes sociales son Twitter o Facebook (Brandwatch, s.f.).

Por otro lado existen APIS para generar el análisis de sentimientos, particularmente usando Excel, dos de los conocidos son Meaningcloud y Lexalytics, tienen características similares, dentro de las cosas que lo diferencian es que en Meaningcloud maneja una gran variedad de idiomas en comparación de la API de Lexalytics, también contiene un número mayor de posibles clasificaciones, las cuales son muy positivo y muy negativo, sin embargo, no permite la modificación del umbral para la clasificación y en los resultados no se muestra el valor numérico (Meaningcloud, s.f.).

Lexalytics tiene a disposición 3 APIS para el estudio de documentos: API en la nube, Semantria for Excel (extensión de Excel) y AI Assembler™, que es una herramienta de modelado de máquinas de aprendizaje (Lexalytics, s.f.).

El software que se utiliza para el análisis de sentimientos en este trabajo será Semantria for Excel, tiene la ventaja de reconocer la importancia de una palabra en un enunciado, evaluando el sentimiento a partir de ellas, es decir que, analiza el contexto de las oraciones en relación al contenido y compara los resultados obtenidos. Por otro lado, permite la modificación total del diccionario de palabras, también posibilita la modificación del umbral para la clasificación, con ello se pueden hacer diferentes experimentos buscando una mejor precisión en la clasificación, es bastante accesible ya que trabaja con Excel y no requiere instancias o programas adicionales utilizando un entorno bastante amigable para operar.

## **3 Análisis de sentimientos en videojuegos**

Como se mencionó, el análisis de sentimientos se enfoca en frases que se clasifican como negativas, positivas o neutrales, se pueden añadir dos evaluaciones más que son, Muy Negativo o Muy Positivo, para el caso de este estudio, se considera cada oración de una conversación como una frase, siendo así un proceso similar a los estudios como SemEval-2016 Task 4. Para lograr

determinar la clasificación de cada conversación, se utiliza un promedio que depende de los valores individuales de las oraciones.

A continuación, se presenta lo realizado en cada etapa de la metodología para realizar el análisis de las conversaciones de los videojuegos LOL, HOTS y OVW.

### **3.1 Recolección y selección**

Para la etapa de recolección de la información, además de las conversaciones se incluyen variables como: el nombre del juego, el idioma, la duración, el tipo de partida, el rango, el resultado obtenido y el género del videojuego, las cuales se explican brevemente a continuación.

- Nombre del juego: se refiere al juego del cual se extrajo la conversación que como se indicó puede ser LOL, OVW y HOTS.
- Idioma: se refiere el idioma predilecto de las conversaciones en la partida.
- Duración: se refiere al tiempo total que duro la partida.
- Rango: en los videojuegos se clasifica a los jugadores en base a su destreza, la clase cambia de juego en juego, por ejemplo, mientras en LOL el nivel más alto es Challenger, en HOTS es Grand master y en OVW es Top 500, para uniformar los datos, en este trabajo se ocuparon los rangos bronce, plata, oro, platino y diamante, siendo bronce el más bajo y diamante el más alto.
- Tipo de Partida: se refiere a diferentes modos de juego, en general se divide en dos, competitiva en donde el resultado de la partida afecta el rango del jugador y casual donde se adquiere experiencia para subir de nivel.
- Resultado Obtenido: se refiere a si la partida concluyó con una victoria o derrota, en LOL se permite la rendición, sin embargo, todos los miembros del equipo deben estar de acuerdo y aún así, se toma como derrota.
- Género: se refiere al tipo de juego, por ejemplo, OVW es un juego FPS, por sus siglas en inglés First Person Shooter, traducido al español como juego de disparos en primera persona y en caso de HOTS y LOL son tipo MOBA por sus siglas Multiplayer Online Battle Arena, si bien, no tiene una traducción exacta, es un juego donde se unen equipos de 5 jugadores para destruir una base enemiga.

Dentro de cada juego, se tiene una interfaz para comunicarse, esta interfaz refleja la conversación de los jugadores de un mismo equipo. Para poder recopilar la información existen dos formas, la primera se logra jugando el videojuego, ya que la partida se almacena junto con las conversaciones en un formato de video. Sin embargo, el video únicamente muestra la conversación entre los integrantes del equipo y no la del equipo rival. La segunda forma es a través de plataformas de contenido multimedia donde los jugadores comparten sus partidas, desafortunadamente, en ocasiones llegan a ser editadas por el jugador por lo que no serían útiles para este trabajo.

En ambos casos, extraer las conversaciones es un proceso tardado ya que requiere, en el caso de que se juegue, el doble del tiempo que dure la partida; por ejemplo, una partida tradicional de LOL dura aproximadamente 45 minutos, más otros 45 minutos para repetirla y capturar el texto. En caso de los videos, solo se requiere acceder a la plataforma y reproducirlos para extraer las conversaciones completas.

Para este trabajo en particular se extrajeron 75 conversaciones, de las cuales, 25 están en idioma español, y 50 en inglés; las conversaciones son de longitud variable, la más extensa contiene 440 palabras, mientras que la más pequeña contiene solo 50 palabras. La longitud de una conversación no influye en su clasificación, pueden existir conversaciones muy largas que sean neutrales y otras que en solo 2 oraciones llegan a ser insultantes o negativas. En lo que respecta a la selección, únicamente se consideraron aquellas conversaciones que tuvieran al menos 50 palabras, como resultado tenemos que en LOL se tienen 3935 palabras que se juntan en 1182 oraciones, para HOTS contiene 1431 palabras con 473 oraciones y para OVW contiene 1681 Palabras con 496 Oraciones, donde cada oración representa a un jugador. Las oraciones tienen entre 1 a 15 palabras en promedio.

Como parte fundamental, la información se presenta de manera anónima, dentro de las conversaciones se vislumbran diferentes patrones de manera inmediata:

- Cada línea es una oración escrita por un jugador, es raro el jugador que escribe más de una línea.
- Se llegan a presentar errores ortográficos y tipográficos.
- En el caso de partidas casuales, la comunicación empieza cuando la partida se empieza a complicar, mientras en las competitivas la comunicación se da desde el inicio.
- Aunque cada jugador tiene un alias, es raro cuando se llegan a mencionar, normalmente se ocupan los nombres de los personajes del juego.
- Especialmente en LOL, para escribir palabras obscenas sin ser censurados, los jugadores optan por cambiar algunas letras por asteriscos.
- Si bien no existen “emojis” en estos juegos, se presentan algunas combinaciones de letras para representarlos.
- Aunque algunas conversaciones son en español, sin embargo, se llega a ocupar palabras o acrónimos en inglés.

Para poder realizar el SA se necesita una organización en las conversaciones, cada videojuego cuenta con algunas palabras o acrónimos específicos, algunas de estas palabras tienen o carecen de un significado en los demás juegos, un ejemplo directo son los personajes de cada juego, un personaje de LOL no existe en HOTS, observando esta característica se han organizado todas las conversaciones con

sus respectivos juegos, de esta manera será más sencillo realizar el diccionario de palabras para cada uno y tener una mejor precisión al realizar el análisis.

### **3.2 Diccionario**

Como se explicó en el punto anterior cada juego se tiene que analizar por separado, sin embargo, así como hay palabras específicas para cada juego, también se tienen palabras con un significado similar sin importar el juego o el idioma, el ejemplo más grande que se tiene es el acrónimo “gg”, que se refiere a las palabras “good game”, y se interpreta como un agradecimiento a los jugadores, es algo similar a un apretón de manos al terminar un partido de fútbol.

Aunque el acrónimo “gg” puede llegar a significar algo positivo, también se puede convertir en algo insultante, si se le añaden las letras “ez”, ya que se convierte en algo totalmente diferente. “ez” hace referencia a la palabra en inglés “easy” debido a su similitud al momento de pronunciarse, “ggez” le da a conocer al enemigo que no fue un rival digno, que las habilidades que mostraron no fueron suficientes para combatir y que fue demasiado fácil alzarse con la victoria, en otras palabras alardear de una supuesta superioridad al equipo enemigo.

Para este trabajo, se agregaron palabras relacionadas con los nombres de los personajes de cada juego, para OVW se añadieron 30 palabras, para HOTS 86 y para LOL 145. Además se insertaron palabras y frases de uso común en los 3 juegos; se añadieron 27 acrónimos como “gg” y por último, se recolectaron todas aquellas palabras o frases que son insultantes o agresivas, entre estas, groserías de uso coloquial, combinación de palabras con doble sentido. También se incluyó una lista negra (“blacklist”) con palabras que incluyen temas como homofobia u odio racial.

Adicional al análisis de sentimientos, Semantria for Excel otorga gráficos en donde representa la frecuencia de las palabras en las conversaciones y su peso informacional en el diccionario, como se aprecia en la tabla 1 y figura 2, existen palabras con un mayor tamaño, esto quiere decir que esa palabra fue mayormente ocupada, el color de la palabra es distintivo de su peso informacional, desde rojo que representa lo negativo, hasta verde que indica palabras positivas, pasando por una escala de grises que son las palabras neutrales que contienen un valor diferente a cero.

Phrase	negative	neutral	positive	Total general
xd	-	23	-	23
gg	-	15	-	15
bien	-	6	1	7
puto	6	-	-	6
vi	-	6	-	6
pendejo	5	-	-	5
bueno	-	5	-	5
:d	-	5	-	5
lux	-	5	-	5
ashe	-	4	-	4
supp	-	4	-	4
mala	4	-	-	4
daño	-	4	-	4
jinx	-	4	-	4
perdón	-	4	-	4
zed	-	4	-	4
buena	-	4	-	4
Total general	15	93	1	109

**Tabla 1.** Frecuencia de las palabras en las conversaciones de LOL

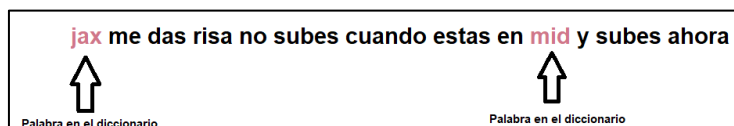


**Figura 2.** Representación gráfica de la frecuencia en tabla 1



### 3.3 Clasificación

Para el análisis de sentimientos, Semantria for Excel busca las oraciones en el diccionario de palabras y añade su valor individual para calcular el valor de toda la conversación y asignarle una clase. Un ejemplo de esto se muestra en la figura 3, en la cual se presenta una oración con las palabras “jax” y “mid” las cuales no parecen tener mucho sentido en la frase.



**Figura 3.** Ejemplo de una oración de LOL

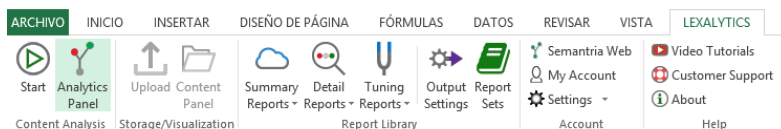
La primera palabra refiere a un personaje del juego “jax” y la segunda a un espacio del mapa como se muestra en la figura 4. En el juego LOL existen 3 líneas donde los personajes pueden transitar para lograr llegar a la base enemiga, la parte de arriba se le conoce como “top”, la de en medio como “mid” (que es la segunda palabra) y la de abajo “bot”; con este conocimiento la oración entonces ya tiene sentido, pues se refiere a que el jugador que maneja al personaje Jax, no trató de llegar a la base enemiga cuando estaba en una posición más sencilla que la actual.



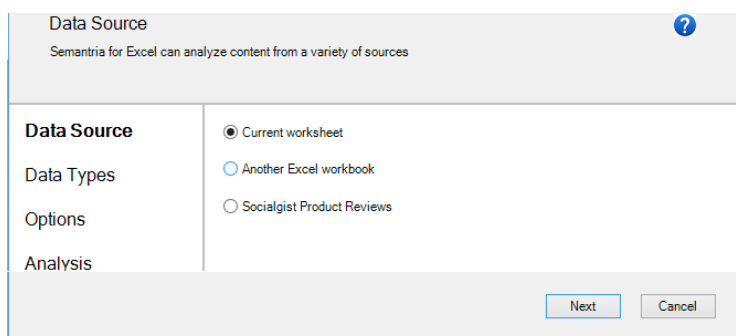
**Figura 4.** Mapa de LOL. Fuente: (Riot, 2009)

Una vez se tengan las conversaciones en un documento de Excel, se puede generar el análisis. En la barra de herramientas de Lexalytics se encuentra un botón el cual lleva por nombre Start (figura 5), al seleccionar esta opción aparecerá una ventana emergente la cual solicita seleccionar el documento con el texto a analizar (figura 6), Semantria solicitará la identificación de las columnas, cuál es el ID, la conversación y los datos extras que pueda contener el documento (figura 7). Posteriormente, se selecciona cuántas filas se van a considerar, pueden ser todas o un intervalo en específico (figura 8), y para concluir, se le da un nombre al análisis, alguna descripción y el diccionario que se ocupará (figura 9).

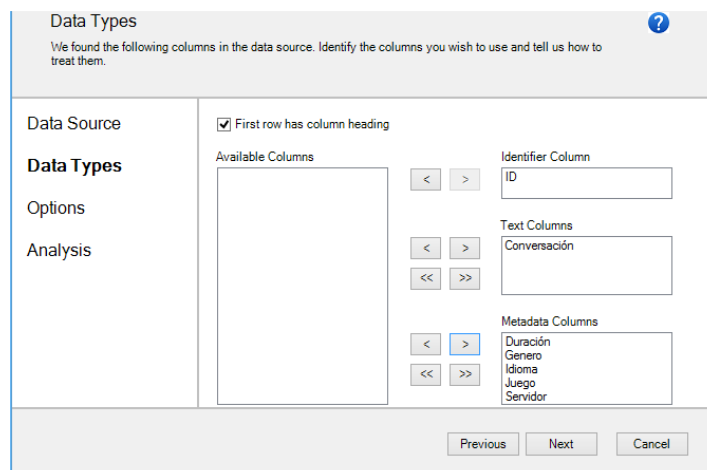
El proceso se inicia y cuando culmine nos mostrará en color verde los procesos realizados (figura 10).



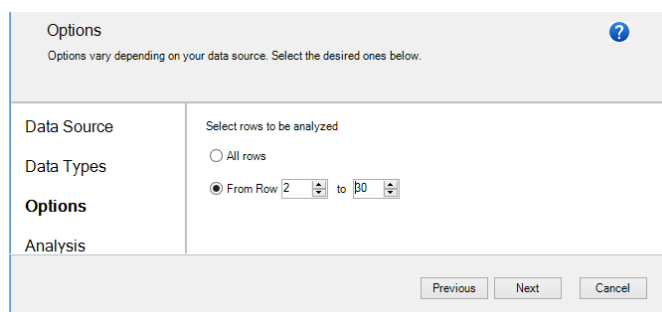
**Figura 5.** Barra de Herramientas de Lexalytics, fuente: (Lexalytics, s.f.)



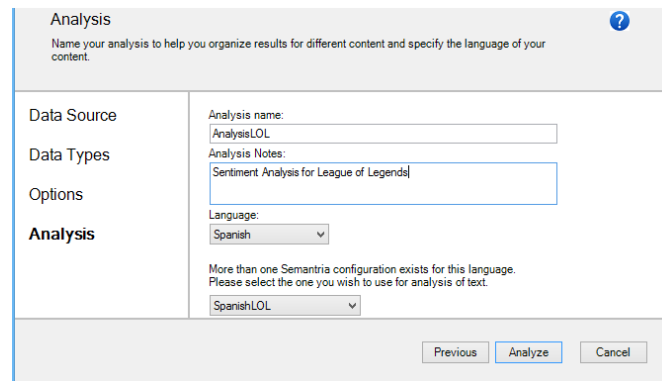
**Figura 6.** Elección de archivo a evaluar, fuente: (Lexalytics, s.f.)



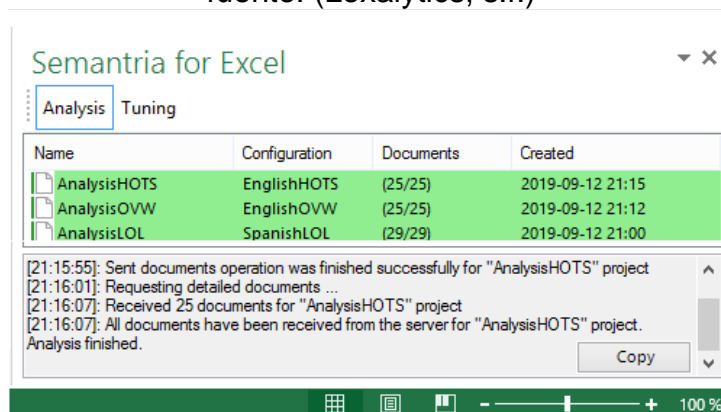
**Figura 7.** Selección de Información fuente: (Lexalytics, s.f.)



**Figura 8.** Tamaño de filas a considerar, fuente: (Lexalytics, s.f.)



**Figura 9.** Nombre, Descripción y Selección del Diccionario para el Análisis, fuente: (Lexalytics, s.f.)



**Figura 10.** Análisis realizados en Semantria, fuente: (Lexalytics, s.f.)

Con base a la documentación proporcionada por Lexalytics para clasificar las conversaciones, Semantria for Excel trabaja de la siguiente forma, se utilizan 3 variables base:

- P = palabra
- F = frase
- N = palabra o frase neutral

Los valores que P o F pueden obtener están en el rango:  $(-\infty, -0.05]$  v  $[0.2, +\infty)$ , mientras que para N, el rango es  $(-0.05, 0.2)$ .

De esta manera, las palabras o frases serán clasificadas como NEGATIVAS, NEUTRALES y POSITIVAS: Si P o F cae en el rango  $(-\infty, -0.05]$  se evalúa como una frase o palabra NEGATIVA, mientras que si cae en el rango  $[0.2, +\infty)$  se evalúa como POSITIVA y finalmente si cae en el rango  $(-0.05, 0.2)$  se evalúa como NEUTRAL.

Para calcular el valor promedio de cada conversación se tiene la ecuación 1:

$$CN = \left( \frac{\sum_{i=1}^{i=n} P_i}{n} + \frac{\sum_{i=1}^{i=n} F_i}{n} + \frac{\sum_{i=1}^{i=n} N_i}{n} \right) / 3 \quad (1)$$

Donde

$P_i$  es el valor de la palabra en el diccionario.

$F_i$  es el valor de la frase en el diccionario

$N_i$  es el valor de las palabras o frases neutrales en el diccionario

Además se deben cumplir las siguientes condiciones:

$$\begin{aligned} P_i &\neq P_{i+1} \\ F_i &\neq F_{i+1} \\ N_i &\neq 0 \end{aligned}$$

En términos generales lo que realiza la fórmula y por consiguiente Semantria for Excel, es realizar la sumatoria de todas las palabras o frases que encuentre en la conversación y al final dividirla entre el número total de palabras o frases. Una vez calculado CN, se calcula TCN, como el valor promedio de todas las conversaciones (ver ecuación 2)

$$TCN = \frac{\sum_{i=1}^{i=NP} CN_i}{NP} \quad (2)$$

Donde:

CN es el valor de la conversación de cada partida y NP el número total de partidas.

El umbral que se utiliza para la clasificación es el recomendado por Semantria for Excel el cual proporciona un punto de partida para determinar si una conversación es positiva o negativa. Dicho umbral se puede modificar con base a las preferencias del usuario y a los patrones que se puedan observar.

### 3.4 Pruebas y resultados

El umbral utilizado para la clasificación de las conversaciones es el dado por defecto en el API de Semantría for Excel. Los resultados se pueden observar en la tabla 2.

Juego	OVW	LOL	HOTS
Conversaciones Positivas	3	1	7
Conversaciones Negativas	11	9	7
Conversaciones Neutrales	11	15	11
Valor promedio de las Conversaciones	-0.0077	-0.0205	0.0032
Promedio de duración	14.28 min	36.9 min	20.45 min
Victorias totales	10	14	7

**Tabla 2.** Resultados del análisis con Semantria for Excel en todas las partidas

En la tabla 2, se aprecia que la cantidad de conversaciones negativas prevalece en los 3 juegos, sin embargo en caso particular de LOL, solo existe 1 positiva; por otro lado, aunque HOTS tiene 11 negativas, es el que tiene la mayor cantidad de conversaciones positivas. En el caso de OVW se observa que existe un equilibrio entre las conversaciones positivas y negativas. Al momento de observar el valor promedio de las conversaciones, que nos refleja que tan Negativas o Positivas fueron las conversaciones se puede ver que aunque LOL no tuvo la mayor cantidad de conversaciones negativas, fue el que tuvo la evaluación más baja, lo que significa que tuvo las conversaciones más negativas. De manera general, en los 3 juegos las conversaciones neutrales son mayoría.

Con el fin de evaluar la certeza del software, se agregó la evaluación de un experto que con base al contenido de las conversaciones las clasificó según su criterio, posteriormente se compararon con los datos ofrecidos por el software de la clasificación dada a la conversación por parte del jugador. De esta forma, la clasificación del sistema obtuvo un 62.6% de certeza, al clasificar de manera correcta 47 de las 75 conversaciones. Para determinar si se podía obtener una mejor precisión en la clasificación, se experimentó utilizando diferentes umbrales para clasificar las conversaciones. En la tabla 3 se muestran los resultados obtenidos.

Umbral	Negativa	Neutral	Positiva	Acierto
<b>-.2 a .2</b>	8	55	12	47
<b>-.5 a .5</b>	3	70	2	41
<b>-.05 a .2</b>	<b>27</b>	<b>37</b>	<b>11</b>	<b>47</b>
<b>-.1 a .1</b>	19	33	23	42
<b>-.05 a .05</b>	27	20	28	36
<b>-.1 a 0.5</b>	19	28	28	37
<b>Experto</b>	<b>18</b>	<b>45</b>	<b>12</b>	<b>75</b>

**Tabla 3.** Resultados de los experimentos

En la tabla 3, se observa que modificar el umbral no aumenta la precisión, el único experimento que dio resultados similares fue el del primer renglón, en el modificar el umbral se incrementó el número de conversaciones neutrales.

#### 4 Aplicación del algoritmo J-48

El algoritmo J-48 tiene una estructura de un árbol de decisión, para lograr realizar la clasificación obtiene un serie de condiciones o reglas con una jerarquía previamente dada, esta jerarquía dependerá del árbol de decisión que se esté ocupando, en base a las variables de entrada devolverá una respuesta para el problema planteado, cada nodo tendrá una salida, estas se dividen en dos partes, las cuales son progresivas o son aquellas que finalizan la clasificación (Aggarwal C., 2015).

En el caso específico del algoritmo J-48, es un algoritmo de inducción, donde se crean subconjuntos de los datos de entrenamiento, en otras palabras, particiones de los datos, en base a estas particiones se generan las reglas pertinentes para lograr clasificar correctamente a el mayor número de casos (Suca, Córdova, Condori, Cayra, & Sulla, 2016).

La precisión del algoritmo va a variar dependiendo del número de particiones que se tengan, aunque en general utiliza el 80% de estos valores como entrenamiento para después probar con el 20% restante, una vez realizada la primer iteración, intercalará entre las particiones, esto quiere decir que hará una combinación de todas las particiones para lograr entrenar y probar entre todas para generar una serie de reglas o condiciones más precisas (Aggarwal C., 2015).

Para aplicar el algoritmo en este caso se añadieron las variables: Nombre del juego, Idioma, Duración de la Partida, Rango del jugador, Tipo de Partida, Resultado Obtenido, Género y la evaluación de las conversaciones, el algoritmo J-48 mostrará, que variable es la más relevante dentro de la clasificación de las

conversaciones, en otras palabras, averiguar si el ambiente (las variables) afecta de manera significativa a la naturaleza de las conversaciones.

En la figura 11, se aprecia los atributos que fueron tomados en cuenta para la aplicación del estudio, las variables que fueron relevantes fueron la puntuación que es la evaluación dada por Semantria for Excel siendo esta variable la indispensable para la clasificación, posteriormente se aprecia que el resultado de la partida infiere en caso de que la evaluación sea muy negativa.

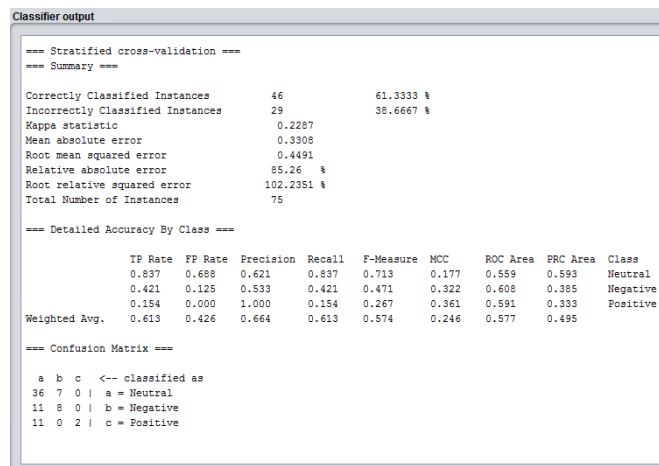
Como se aprecia en la figura 12 la precisión que muestra el algoritmo es de 61.33%, en comparación de la dada por Semantria for Excel no se observa un aumento, esto debido que el algoritmo toma como variable principal las conversaciones, adicional, se muestra la matriz de confusión del algoritmo, el comportamiento que ofrece es el clasificar en gran medida las conversaciones en Neutral debido a la gran cantidad que existen, en la figura 13 se observa el árbol del algoritmo mostrando de manera gráfica este comportamiento.

```
Classifier output
=== Run information ===
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    datos generales-weka.filters.unsupervised.attribute.Remove-R1
Instances:   75
Attributes:  8
             Juego
             Punt
             Duración
             Resultado
             Tipo de Partida
             Genero
             Rank
             Clasificación
Test mode:   5-fold cross-validation

=== Classifier model (full training set) ===

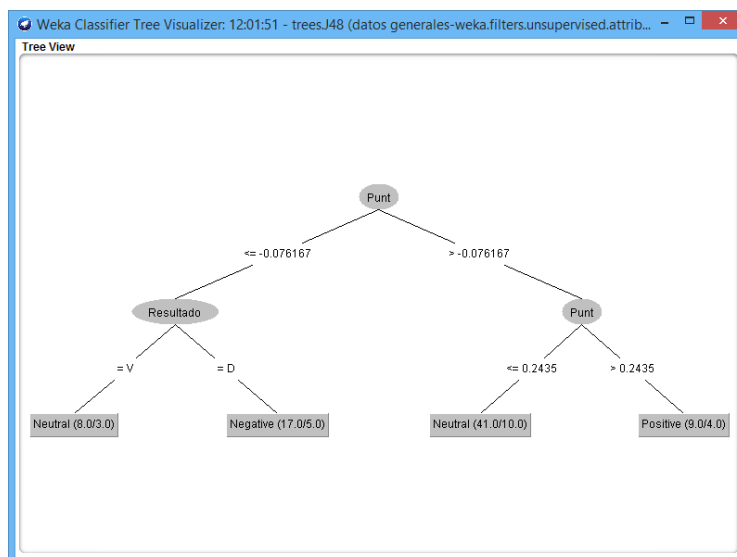
J48 pruned tree
-----
Punt <= -0.076167
| Resultado = V: Neutral (8.0/3.0)
| Resultado = D: Negative (17.0/5.0)
Punt > -0.076167
| Punt <= 0.2435: Neutral (41.0/10.0)
| Punt > 0.2435: Positive (9.0/4.0)
```

Figura 11. Atributos y comportamiento del algoritmo J-48



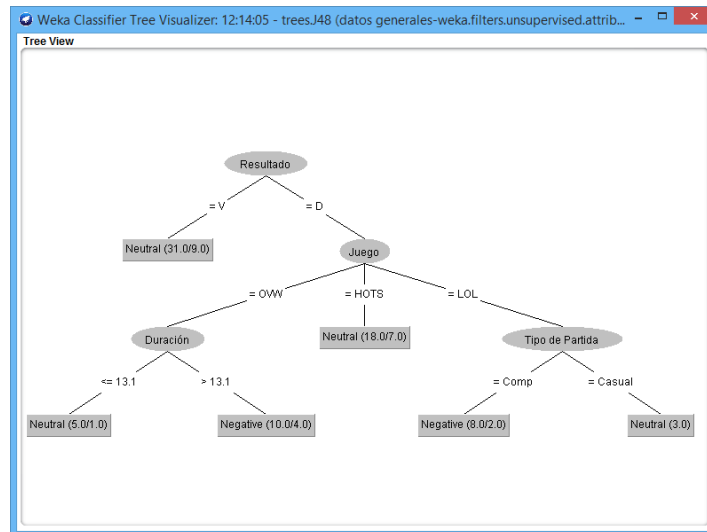
**Figura 12.** Precisión en la clasificación y matriz de confusión del algoritmo J-48

Al excluir la variable de puntuación (Punt), que es la dada por Semantria, observamos un comportamiento en donde las variables principales que se toman en consideración son el resultado del juego (Victoria o Derrota), en caso de derrota verifica el juego (OVW, HOTS o LOL) y concluye con revisar el tipo de partida o la duración, este comportamiento se puede apreciar en la figura 14.



**Figura 13.** Árbol de clasificación J-48





**Figura 14.** Comportamiento del algoritmo sin la variable puntuación

## 5 Conclusiones y trabajo futuro

En este artículo se presentaron los resultados de aplicar el análisis de sentimientos a 75 conversaciones de longitud variable escritas en inglés y español, extraídas de partidas de los videojuegos *League of Legends*, *Overwatch* y *Heroes of the Storm*. Utilizando Semantría para Excel, un API del software Lexalytics, se creó un diccionario que pudiera ser utilizado para la clasificación de las conversaciones. Las clases se definieron como positivas, negativas o neutrales.

Por medio de los experimentos realizados se observa que sí es posible aplicar el análisis de sentimientos en las conversaciones de una partida de videojuegos, obteniendo un 62.6% de certeza en la clasificación. Aunque el resultado puede considerarse bajo, se justifica por el desbalance en el número de conversaciones de cada clase, es necesario tener una cantidad equitativa de los tres tipos de conversaciones (Negativa, Positiva y Neutral) esperando obtener un mayor grado de certeza.

Si bien existe un desbalance en las conversaciones, se atribuye a que en gran medida las conversaciones tienden a ser negativas reflejando un comportamiento dentro de las 3 comunidades (*Overwatch*, *League of Legends* y *Heroes of the Storm*), en general se aprecia que así como en otros deportes, las riñas llegan a presentarse en cualquier nivel de juego, siguiendo el mismo concepto que se ha manejado siempre, la competitividad.

Por otro lado el algoritmo C4.5 o J-48 mostró la naturaleza del comportamiento de las clasificaciones y ratificó que en esencia los implicados en las conversaciones tienen mucho que ver en sus comportamientos Negativos o Positivos.

Como trabajo futuro, se considera contar con un mayor número de conversaciones, resolver el problema del desbalance, así como extender de tres a cinco, las clases para mejorar la clasificación incluyendo la actualización del diccionario.

Se considera hacer un análisis comparativo de algoritmos de clasificación con la finalidad de observar si se tienen mejores resultados que el algoritmo C4.5 que permita aumentar la certeza de la clasificación.

## REFERENCIAS

- Aggarwal C., C. (2015). *Data Mining*. Switzerland: Springer. doi:10.1007/978-1-4614-3223-4
- Brandwatch. (s.f.). *Brandwatch Analytics*. Recuperado el 16 de 04 de 2019, de <https://www.brandwatch.com/es/products/analytics/>
- Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, 2, 14. doi: 10.1186/s40537-015-0015-2
- Harris, B. (2014). *Console Wars: Sega Vs Nintendo-and the Battle that Defined a Generation*. Atlantic Books Ltd.
- Lexalytics. (s.f.). *Data Analytics with NLP & Text Analytics*. Recuperado el 20 de 04 de 2019, de <https://www.lexalytics.com/technology/summarization>
- Mat Zin, H., Mustapha, N., Azmi Murad, M., & Mohd Sharef, N. (2017). The Effects of Pre-Processing Strategies in Sentiment. *AIP conference proceedings*, 20-89. doi:10.1063/1.5005422
- Meaningcloud. (s.f.). *Meaningcloud Sentiment Analysis in Excel*. Recuperado el 03 de 04 de 2019, de <https://www.meaningcloud.com/developer/excel-addin/doc/3.4/sentiment-analysis>
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications. A survey. *Ain Shams engineering journal*, 5, 1093-1113. doi:10.1016/j.asej.2014.04.011
- Navok, P., Ritter, A., Rosenthal, S., Sebastiani, F., & Stoyanov, V. (2016). SemEval-2016 Task 4: Sentiment Analysis in Twitter. *Proceedings of the 10th international workshop on semantic evaluation (semeval-2016)*, 1-18. doi:10.18653/v1/S16-1
- Riot, G. (27 de Octubre de 2009). League of Legends. USA. Obtenido de <https://lan.leagueoflegends.com/es/game-info/>

Rosenthal, S., Farra, N., & Nakov, P. (2017). SemEval-2017 Task 4: Sentiment Analysis in Twitter. *Proceedings of the 11th international workshop on semantic evaluation*, 502-518. doi:10.18653/v1/S17-2088

Steven, K. (2016). *The Ultimate History of VideoGames*. B DE BOOKS.

Strååt, B., & Verhagen, H. (2017). Using User Created Game Reviews for Sentiment. *GHITALY@CHIItaly*.

Suca, C., Córdova, A., Condori, A., Cayra, J., & Sulla, J. (2016). Comparación de algoritmos de clasificación para predicción de casos de obesidad infantil. *Perú: Universidad Nacional de San Agustín*.

Waikato, U. d. (31 de 10 de 2011). Weka. Nueva Zelanda: GNU-GPL.

## Notas biográficas



**Daniel Iván Ángeles Gómez** Recibió el certificado total de estudios de la carrera de Ingeniería en Sistemas y Comunicaciones en la Universidad Autónoma del Estado de México. Actualmente, se encuentra en la búsqueda de su titulación y crecimiento profesional. Sus áreas de interés son el desarrollo de software, análisis de datos y desarrollo de videojuegos.



**Dra. Maricela Quintana López** es Profesora de Tiempo Completo del Centro Universitario UAEM Valle de México de la Universidad Autónoma del Estado de México. La Dra. Quintana es Ingeniera en Sistemas Computacionales, y tiene la Maestría y el Doctorado en Ciencias Computacionales con especialidad en Inteligencia Artificial por el Tecnológico de Monterrey. Cuenta con el reconocimiento de Perfil Deseable PRODEP de la SEP en México. Ha dirigido tesis de maestría y licenciatura, ha sido conferencista magistral y ponente en diversos eventos. Ha publicado artículos en las áreas de Inteligencia Artificial y Minería de Datos relacionados con sus proyectos de Investigación.



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 2.5 México.

*Recibido 01/06/2019*

*ReCIBE, Año 8 No. 2, Noviembre 2019*

*Aceptado 23/08/2019*

## **Aplicación de Algoritmos de Votación en la Clasificación de Textos**

**Saturnino Job Morales Escobar<sup>1</sup>**

sjmoralese@uaemex.mx

**María Esther Guevara Cruz<sup>2</sup>**

ttguevara@yahoo.com.mx

**Héctor Rafael Orozco Aguirre<sup>1</sup>**

hrorozcoa@uaemex.mx

<sup>1</sup>Centro Universitario UAEM Valle de México  
Universidad Autónoma del Estado de México

<sup>2</sup>Universidad Tecnológica Fidel Velázquez

**Resumen:** Los algoritmos de clasificación basada en la similitud entre textos, es un tópico de gran interés lo que es motivado por la relevancia y complejidad que presenta esta problemática y la gran cantidad de ámbitos que requieren de sistemas automatizados para resolverla. Por mencionar algunos ejemplos, se requiere de aplicaciones capaces de determinar de manera automática la sensibilidad de la información textual almacenada en archivos para decidir el grado de protección o diferenciar datos sensibles de aquellos que no lo son para prevenir o detectar su fuga o accesos no autorizados. Esto puede ser establecido con base en la importancia de su contenido. Así, una forma de abordar el problema es desde la perspectiva de la clasificación de textos con base en contexto. Para ello, es necesario desarrollar algoritmos que consideren la semántica de los documentos desde el momento que se generan y obtener su clasificación mediante un conjunto de características que incluyan rasgos textuales. En este trabajo, se definen los rasgos textuales y la relación semántica entre rasgos textuales. Considerando que un texto es la concatenación de rasgos textuales, la relación semántica de los rasgos textuales permite determinar su grado de pertenencia a un tema. Para la clasificación de textos, se aplican las etapas del modelo de algoritmos de votación desarrollados bajo el Reconocimiento Lógico Combinatorio de Patrones.

**Palabras clave:** clasificación de textos, algoritmos de votación, rasgos textuales, contexto.

**Abstract:** Classification algorithms based on the similarity between texts, is a topic of great interest which is motivated by the relevance and complexity of this problem and the large number of areas that require automated systems to solve it. To mention some examples, applications capable of automatically determining the sensitivity of textual information stored in files are required to decide the degree of protection or differentiate sensitive data from those that are not to prevent or detect their leakage or unauthorized access. This can be established based on the importance of their content. Thus, one way to approach the problem is from the text classification perspective based on context. For this, it is necessary to develop algorithms that consider the document semantics from the moment they are generated and obtain their classification through a set of features that include textual features. This paper defines the textual features and the semantic relationship between textual features. Considering that a text is the concatenation of textual features, the semantic relationship of textual features makes possible to determine their degree of belonging to a topic. For text classification, the stages of the voting algorithm model developed under the Combinatory Logical Pattern Recognition are applied.

**Keywords:** Text classification, voting algorithms, textual features, context.

## 1. Introducción

Como consecuencia del gran incremento en la cantidad de información que se genera en diversos ámbitos del quehacer humano, se incrementa también la necesidad de herramientas que ayuden en la actividad humana de clasificarla, ya que hacerlo de manera manual, es prácticamente imposible. Como caso particular de la información que se genera en formatos digitales, se podría considerar la contenida en documentos textuales. En este caso, la clasificación de documentos está relacionada a diversos fines, los cuales van desde su recuperación en grandes almacenes de información, hasta la determinación de la importancia de su contenido para implementar mecanismos de protección para evitar que usuarios no autorizados tengan acceso a ella. Esto ha motivado que la clasificación de textos sea estudiada desde diferentes enfoques y en diversas áreas de investigación de las Ciencias de la Computación y la Informática, desarrollando desde medidas de similitud, según trabajos como los dados en (Kumar & Basha, 2015; Coello, y otros, 2018), hasta diversos algoritmos de clasificación desde varios enfoques.

Desde el punto de vista metodológico, para abordar el problema, es necesario llevar a cabo varias etapas por medio de las cuales se puedan formalizar los criterios y reglas que se utilizan como base para la clasificación de textos. Es decir, se requiere que, en una primera etapa, se definan las clases en las cuales se pueden ubicar los textos, se determinen las variables mediante las cuales se describirán los textos, los conjuntos de valores asociados a cada variable y cómo se efectuará la comparación de esos valores, en otra etapa, formalizar las relaciones entre términos y temas, el criterio para determinar cómo comparar los textos, hasta finalmente, precisar el criterio para su clasificación. En la solución del problema de la clasificación de textos, existe una gran diversidad de métodos y algoritmos, los cuales cubren desde la extracción de características textuales y la reducción de características, hasta la clasificación, los cuales están basados en árboles de decisión, redes neuronales y reglas estadísticas. Para consultar un estudio reciente y muy completo se puede ver lo hecho en el trabajo presentado en (Kowsari, y otros, 2019) y lo realizado en (Aggarwal & Zhai, 2012). Además, existen otros enfoques para resolver el problema, como el Procesamiento del Lenguaje Natural (PLN), considerado una subárea de la Inteligencia Artificial y la Lingüística, con el objetivo de estudiar los problemas derivados de la generación y comprensión automática del lenguaje natural.

El procesamiento estadístico del lenguaje natural, Manning & Schütze (1999), representan el modelo clásico de los sistemas de recuperación de información, este modelo se caracteriza porque cada documento está descrito por un conjunto de palabras clave denominadas términos índices. Este enfoque se basa en lo que se ha denominado como bolsa de palabras. En esta aproximación, todas las palabras de un documento se tratan como términos índices para ese documento. Además, se asigna un peso a cada término en función de su importancia, este peso es determinado normalmente con base en su frecuencia de aparición en el documento. De este modo, no se toma en consideración el orden, la estructura, el significado, etc. de las palabras según Vallez & Pedraza (2007).

En otros enfoques, se han incorporado como herramienta la Teoría de Subconjuntos Difusos y la Lógica Difusa para modelar la imprecisión y la incertidumbre. En este trabajo, la clasificación de textos se aborda como un problema de clasificación supervisada desde el enfoque del Reconocimiento Lógico Combinatorio de Patrones (RLCP). En este enfoque, se requiere: determinar los rasgos a utilizar para la descripción de los textos; definir el criterio con el cual se compararán los textos para determinar si son similares o no, o que tan similares son; determinar las clases y elegir las muestras de entrenamiento y de control (para evaluar la eficiencia del clasificador) y elegir el algoritmo o familia de algoritmos a utilizar. Entonces el proceso de crear una clasificación consiste en evaluar las relaciones entre los rasgos de los objetos ya clasificados con los del objeto a clasificar y decidir la relación de pertenencia del objeto a las distintas clases.

En este trabajo, se reporta la aplicación del modelo de algoritmos de votación (ALVOT), y la determinación de las relaciones semánticas en la clasificación de trabajos finales para obtener el título correspondiente en algunas carreras del Centro Universitario UAEM Valle de México de la Universidad Autónoma del Estado de México, México. Los documentos se describieron utilizando las variables o rasgos: título, resumen, carrera y asesor, y se excluyeron, por ejemplo, fecha de terminación, opción de titulación por considerar que no se influyen en la clasificación. Estos textos o documentos fueron clasificados por área de conocimiento de acuerdo al tema. En una primera instancia, el conjunto de temas es definido por el especialista, en este caso, por los autores del trabajo y sus asesores.

## 2. Fundamentos teóricos

La representación de los objetos a clasificar, en este trabajo textos, es un aspecto importante que se aborda en el proceso de formalización de la solución de un problema en el RLCP (Martínez-Trinidad & Guzmán-Arenas, 2001), los objetos se describen mediante un conjunto de variables que da lugar a un espacio de representación formado por el producto cartesiano de los dominios de dichas variables. Así, un objeto se representa por un n-uplo donde se pueden mezclar variables tanto cuantitativas como cualitativas. Para el planteamiento formal y la notación utilizada en el presente documento, se toma la dada en (Ruiz-Shulcloper, Arenas, & Trinidad, 1999).

Sea  $U$  un universo de objetos estructurados en  $l$  clases  $K_1, K_2, \dots, K_l$ . Los objetos están descritos en términos del conjunto de variables  $R = \{x_1, x_2, \dots, x_n\}$ , para cada elemento  $O_i \in U$  la descripción está dada por el n-uplo  $I(O_i) = (x_1(O_i), \dots, x_n(O_i))$ .



Sea  $M_p$  el conjunto de valores admisibles de la variable  $x_p$ ,  $p = 1, \dots, n$ , es decir:

$$x_p(O_i) \in M_p \quad i = 1, \dots, m. \quad (1)$$

Para cada objeto  $O_i$  de  $U$  hay una  $l$ -upla de pertenencia:

$$\alpha(O_i) = (\alpha_1(O_i), \dots, \alpha_l(O_i)) \text{ donde } \alpha_j(O_i) \in \{0,1\}, \quad (2)$$

El valor 1 indica que el objeto pertenece a la clase, 0 que no pertenece.  $K_j$ ,  $j = 1, \dots, l$ . Esta información se almacena en una estructura denominada matriz de entrenamiento (ME).

En la Tabla 1, se muestra la estructura del contenido de la matriz de entrenamiento, donde  $x_i(O_j)$  representa el valor de la variable  $x_i$  en el objeto  $O_j$ .

	$x_1$	...	$x_n$	$\alpha_1$	...	$\alpha_l$
$O_1$	$x_1(O_1)$	...	$x_n(O_1)$	$\alpha_1(O_1)$	...	$\alpha_l(O_1)$
.	.	...	.	.	...	.
.	.		.	.		.
.	.		.	.		.
$O_m$	$x_1(O_m)$	...	$x_n(O_m)$	$\alpha_1(O_m)$	...	$\alpha_l(O_m)$

**Tabla 1.** Matriz de entrenamiento (ME)

**Definición 2.1.** Un criterio de comparación para la variable  $x_i \in R$  es una función  $C_i: M_i \times M_i \rightarrow V_i$ , en donde  $\forall i, V_i$  es un conjunto totalmente ordenado. Esta función da el grado de similitud entre un par de valores admisibles para  $x_i$ .  $C_i$  puede ser booleano, k-valente o real.

**Definición 2.2.** Sea  $M_{i_1} \times M_{i_2} \times \dots \times M_{i_s}$  para cualquier  $T = \{x_{i_1}, x_{i_2}, \dots, x_{i_s}\} \subseteq R$ , siendo  $R$  el conjunto de variables en términos en los que son descritos los objetos, se define una función  $\beta: (M_{i_1} \times M_{i_2} \times \dots \times M_{i_s})^2 \rightarrow V$ , donde  $V$  es un conjunto totalmente ordenado y  $\beta$  es una función de semejanza (parcial) con denominaciones análogas a  $C_i$  en dependencia de  $V$ , la cual cumple  $\beta_T(O, O) \geq \max_{O' \in M} \{\beta_T(O, O')\}$ . Cuando  $s = n$ , se dice que  $\beta$  es una función de semejanza total.

**Definición 2.3.** Sea  $L$  un lenguaje definido sobre un alfabeto  $\Sigma$ ,  $t \in L$  un término y  $C$  un tema. Se define la relación semántica entre  $t$  y  $C$  como el grado de asociación entre  $t$  y  $C$ , el cual se denota como  $\mu_C(t)$ , donde  $\mu_C(t) \in [0,1]$ . Esta asociación se puede ver como el grado de pertenencia de un término a un tema.

**Definición 2.4.** Sea  $t_x = t_1 t_2 \dots t_s$  la concatenación de términos para formar un texto, la semejanza semántica de  $t_x$  con  $C$  se define como el grado de asociación entre  $t_x$  y  $C$ , y se denota por  $\mu_C(t_x)$ , donde  $\mu_C(t_x) \in [0,1]$ . Aquí, la semejanza semántica proporciona la relación entre el texto y el tema. En este caso, el grado de asociación, se calcula sumando los grados de asociación de cada término al tema correspondiente.

**Definición 2.5.** Un rasgo textual es aquel cuyo conjunto de valores admisibles está definido como una concatenación de términos.

### 3. Formalización del problema de clasificación

Desde el enfoque del RLCP, una manera de resolver problemas de clasificación supervisada es a través del modelo de ALVOT, lo que implica que para definir un ALVOT se debe cumplir cada una de las siguientes etapas de acuerdo a lo dado en (Ruiz-Shulcloper, Arenas, & Trinidad, 1999):

1. Determinar el sistema de conjuntos de apoyo.
2. Determinar la función de semejanza.
3. Evaluación por fila para un conjunto de apoyo.
4. Evaluación por clase para un conjunto de apoyo.
5. Evaluación por clase para el sistema de conjuntos de apoyo.
6. Regla de solución.

Con base en este modelo, en cada etapa se deben definir los parámetros para determinar un algoritmo, a continuación, se describe cada etapa:

1. La determinación del sistema de conjuntos de apoyo consiste en elegir subconjuntos de rasgos ( $\Omega_i \subseteq R$ ), para formar un conjunto  $\Omega$  (sistema de conjuntos de apoyo). Los conjuntos de apoyo serán utilizados en la función de semejanza para la comparación de las subdescripciones de objetos. Como posibles sistemas de conjunto de apoyo se pueden tomar, desde el conjunto potencia de rasgos, hasta conjuntos dados por el especialista.
2. En la determinación de la función de semejanza, se puede utilizar alguna función ya sea conocida, o bien, definir una nueva función que refleje el criterio sobre el cual se basará la comparación entre pares de objetos. En el caso del presente trabajo, esto se logró con la definición de semejanza semántica.
3. La evaluación por fila para un conjunto de apoyo permite expresar, por medio de una fórmula, cómo se compara cualquier objeto de la muestra de entrenamiento con el objeto a clasificar. En la fórmula se puede incluir, una ponderación de la semejanza con base en la importancia del objeto (documento) de la ME utilizado en la comparación y las ponderaciones correspondientes a la importancia informacional o relevancia de los rasgos comparados (conjunto de apoyo).
4. La evaluación por clase para un conjunto de apoyo permite totalizar las evaluaciones obtenidas por el documento a clasificar en cada clase y con respecto a cada conjunto de apoyo.
5. En la etapa de evaluación por clase para el sistema de conjuntos de apoyo, se totaliza la evaluación para todos los conjuntos de apoyo. En esta etapa, se tiene la posibilidad de ponderar cada conjunto de apoyo de manera diferenciada e integrar las evaluaciones parciales de la comparación del documento a clasificar con respecto a cada subconjunto de rasgos.

6. Finalmente, en la determinación de la regla de solución, se expresa el criterio con base en el cual se toma la decisión con relación a la pertenencia del documento a cada una de las clases. La decisión de pertenencia nuevamente puede ser booleana, k-valente o real y es expresada por medio del l-uplo dado como respuesta del algoritmo.

#### 4. Resultados

Con base en lo presentado en la sección previa, para la aplicación de la propuesta de clasificación de textos se utilizó como caso de estudio un conjunto de trabajos escritos para titulación clasificados por tema. Así, el universo fue un conjunto con 166 trabajos de titulación clasificados en las siguientes áreas: tratamiento de la información, redes, ingeniería de software, programación, minería de datos, software de base, inteligencia artificial y hardware. Estos trabajos de titulación han sido realizados por alumnos egresados del Centro Universitario UAEM Valle de México de la Universidad Autónoma del Estado de México, México. Los documentos están descritos en términos del conjunto de rasgos  $R = \{\text{titulo}, \text{resumen}, \text{asesor}, \text{carrera}\}$ . Así, la descripción de cada documento será determinada por el título del trabajo, el resumen, el asesor y la carrera cursada por el estudiante. De esta manera, la descripción de un documento será una fila de la matriz de entrenamiento o de la matriz de control, la cual tiene la misma estructura que la matriz de entrenamiento y es utilizada para evaluar la eficiencia del algoritmo de clasificación.

Como se puede observar de las descripciones, todos los rasgos son de tipo textual y los conjuntos de los valores admisibles para cada rasgo están determinados como un conjunto de rasgos textuales. En este caso, como primera aproximación, se consideró que cada documento fuera clasificado solamente en un tema, sin embargo, se puede eliminar esta restricción. La matriz de entrenamiento  $ME = \{d_1, d_2, \dots, d_n\}$ , se conformó con 94 trabajos de carrera afines, los restantes 72 se utilizaron para la matriz de control, que como se mencionó, será utilizada para evaluar la calidad de la clasificación de los documentos.

Es importante notar, que como en todos los problemas que involucran muestra de entrenamiento, los resultados dependen fuertemente de la selección de esa muestra de entrenamiento. De las pruebas realizadas, los conjuntos de apoyo fueron tomados del conjunto potencia de rasgos, sin establecer un orden en la elección. Como se sabe, el total de subconjuntos es 16, pero el conjunto vacío no es considerado. A continuación, se muestran algunos ejemplos de conjuntos de apoyo:

$$\begin{aligned} \Omega_1 &= \{\text{titulo}, \text{resumen}, \text{asesor}\}, \Omega_2 = \{\text{resumen}\}, \Omega_3 = \{\text{titulo}, \text{carrera}\}, \\ \Omega_4 &= \{\text{resumen}\}, \Omega_5 = \{\text{titulo}, \text{resumen}, \text{carrera}\}, \Omega_6 = \{\text{titulo}\}, \\ \Omega_9 &= \{\text{titulo}, \text{resumen}\}, \Omega_{16} = \{\text{titulo}, \text{resumen}, \text{asesor}, \text{carrera}\} \end{aligned}$$

Los resultados sobre las pruebas de clasificación que se presentan en este trabajo fueron seleccionados tomando los conjuntos de apoyo  $\Omega_2$ ,  $\Omega_6$  y  $\Omega_9$ , los cuales aportan una mayor cantidad de información. Con base en los temas definidos para la clasificación y los términos más frecuentemente utilizados en esa área, se determinó la relación semántica entre los términos y cada tema, los cuales forman la matriz de asociación semántica. En la Tabla 2, se muestran ejemplos de las asociaciones semánticas obtenidas.

Término	Tema							
	Tratamiento de la Información	Redes	Ingeniería de Software	Programación	Minería de Datos	Software Base	Inteligencia Artificial	Hardware
<b>Red</b>	0.6	1	0.7	0.5	0.8	0.1	0.65	0.3
<b>Servidor</b>	0.2	0.5	0.4	0.15	0.3	0.6	0.1	0.45
<b>Software</b>	0.3	0.2	0.6	0.4	0.5	0.55	0.45	0.1
<b>Organización</b>	0.7	0.5	0.4	0.15	0.45	0.55	0.1	0.12

**Tabla 2.** Relación semántica entre temas y términos

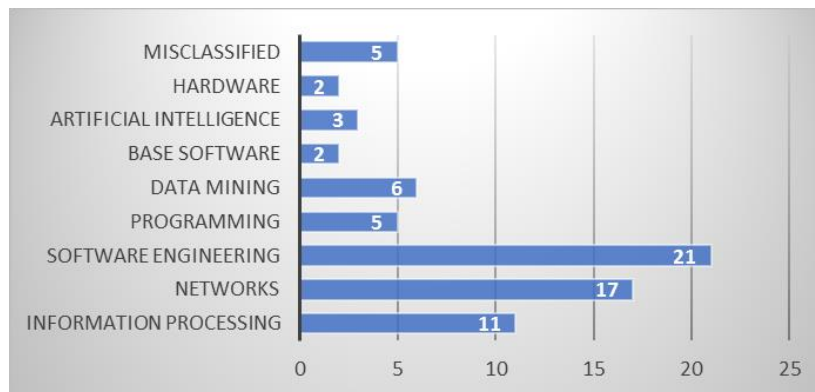
De manera análoga, se determinó una matriz de la semejanza semántica para los valores de los rasgos carrera y asesor. En la Tabla 3, se muestra un ejemplo de las descripciones de los documentos en términos de los rasgos almacenados en la matriz de entrenamiento.

A cada documento, se le asocia un peso informacional relativo a su clase, el cual depende de la representatividad del documento con respecto al tema. La función de semejanza se determinó como el valor promedio entre las semejanzas semánticas de cada rasgo textual. La decisión sobre la pertenencia o no a una clase se tomó con base en las evaluaciones, así, un documento se ubica en la categoría en la cual obtenga la evaluación máxima.

Título	Resumen	Carrera	Asesor
Clasificación de series de tiempo mediante ...	La clasificación de series de tiempo ha atraído gran interés en diversas áreas como economía, medicina, industria, entre muchas otras, en las cuales se hace necesario una clasificación para poder estudiarlas ...	Ingeniería en Sistemas y Comunicaciones	Saturnino Job Morales Escobar

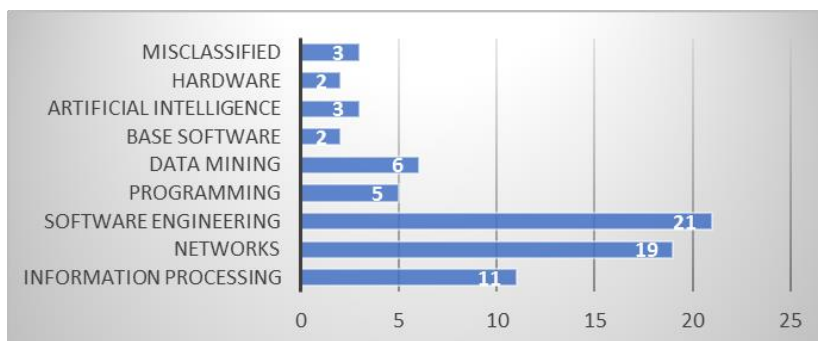
**Tabla 3.** Ejemplo de la descripción de un documento en términos de los rasgos utilizados

Como resultado, se tiene un programa que realiza clasificación de textos con base en el modelo de ALVOT, la relación semántica de términos y la semejanza semántica. Se hicieron corridas considerando 72 documentos en la matriz de control y diferentes conjuntos de apoyo. A continuación, se presentan los resultados con los conjuntos de apoyo  $\Omega_2$ ,  $\Omega_6$  y  $\Omega_9$ . Considerando la función de semejanza sobre el conjunto de apoyo  $\Omega_6$  se obtuvieron los resultados que se muestran en la gráfica de la Tabla 4.



**Tabla 4.** Clasificación de documentos por área de conocimiento con base en el título

Ahora, cuando la función de semejanza se aplicó sobre el conjunto de apoyo  $\Omega_2$  se obtuvieron los resultados que se muestran en la gráfica de la Tabla 5. Obteniendo una mejor clasificación de acuerdo con la matriz de control.



**Tabla 5.** Clasificación de documentos por área de conocimiento con base en el resumen

En la tercera prueba, se tomaron en cuenta los rasgos título y resumen ( $\Omega_9$ ) del documento, en esta clasificación de los 72 documentos se clasificaron correctamente 70, como se puede apreciar en la Tabla 6.



**Tabla 6.** Clasificación de documentos por área de conocimiento con base en título y resumen.

De los resultados mostrados en las tablas anteriores, cuando se utilizó solamente el título, la eficiencia de la clasificación fue del 95%, al agregar el resumen subió al 96% y al utilizar ambos rasgos se logró el 98%, lo que representa una buena calificación para el clasificador.

Clasificación	Con base en el título	Con base en el resumen	Con título y resumen
<b>Eficiencia</b>	93%	96%	98%

**Tabla 7.** Eficiencia de la clasificación para los conjuntos de apoyo  $\Omega_2$ ,  $\Omega_6$  y  $\Omega_9$

## **5. Conclusiones y trabajo futuro**

En los resultados se observó que, en algunas de las corridas, la clasificación basada solamente en el título del documento producía una clasificación errónea, pero al realizar la prueba considerando también el resumen, se mejoró la clasificación. La causa es que el título del documento no siempre ofrece información suficiente sobre lo que en realidad tratará el mismo, sin embargo, el resumen proporciona un panorama más amplio del tema del documento.

De la misma manera, de los resultados obtenidos, se puede observar que la inclusión de nuevos tipos de rasgos, las relaciones y semejanzas semánticas en el modelo de ALVOT en el proceso de clasificación ofrece una solución a problemas de clasificación de documentos considerando el contexto. No obstante, los resultados de las pruebas realizadas, es importante continuar con los trabajos en este campo, de tal manera que la obtención de las relaciones y semejanzas semánticas sean obtenidas de forma sistematizada de acuerdo con otros modelos. Con base en los resultados obtenidos, se puede aplicar el modelo de ALVOT y las relaciones semánticas en la determinación de la sensibilidad de documentos textuales para establecer con ello mecanismos de protección adecuados.

Finalmente, con lo realizado en el trabajo reportado en el presente documento, se avanza en el tratamiento de textos considerando el contexto, de tal manera que es posible continuar los trabajos orientando los resultados hacia la comprensión de lenguaje natural y el procesamiento automático de información que tiene aplicaciones a nivel global de la Inteligencia Artificial, en particular, a aplicaciones en Inteligencia de Negocios.

## REFERENCIAS

- Aggarwal, C. C., & Zhai, C. (2012). A survey of text clustering algorithms. En C. C. Aggarwal, & C. Zhai, *Mining text data* (págs. 77-128). Boston, MA, USA: Springer.
- Coello, L., Frias, M., Fernandez, Y., Filiberto, Y., Bello, R., & Caballero, Y. (2018). Construcción de relaciones de similaridad borrosa basada en la medida calidad de la similaridad. *Investigación Operacional*, 38(2), 132-140.
- Kowsari, K., Jafari Meimandi, K., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text Classification Algorithms: A Survey. *Information*, 10(4), 1-68. Obtenido de Information: <https://arxiv.org/pdf/1904.08067.pdf>
- Kumar, B. V., & Basha, M. S. (2015). Optimal Similarity Measure to Ensure Robustness in Text Classification and Clustering. *International Journal of Electronics Communication and Computer Engineering*, 6(5), 161-167.
- Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. London, England: MIT Press.
- Martínez-Trinidad, J. F., & Guzmán-Arenas, A. (2001). The logical combinatorial approach to pattern recognition, an overview through selected works. *Pattern Recognition*, 34(4), 741-751.
- Ruiz-Shulcloper, J., Arenas, A. G., & Trinidad, J. F. (1999). *Enfoque Lógico Combinatorio al Reconocimiento de Patrones*. Ciudad de México, México: Instituto Politécnico Nacional.
- Vallez, M., & Pedraza, R. (20 de september de 2007). El Procesamiento del Lenguaje Natural en la Recuperación de Información Textual y áreas afines. *Hipertext.net*(5). Obtenido de Hipertext.net: <http://www.hipertext.net>



## Notas biográficas



**Saturnino Job Morales Escobar** recibió el título de Licenciado en Computación y el grado de Maestro en Ciencias de la Computación por la Benemérita Universidad Autónoma de Puebla, México. Ha participado en el desarrollo de proyectos de investigación en las áreas de la Inteligencia Artificial y el Reconocimiento de Patrones y sus aplicaciones.

Desde 2002 es Profesor de Tiempo Completo en el Centro Universitario Valle de México de la Universidad Autónoma del Estado de México, México. Pertenece al cuerpo académico de Inteligencia Computacional y sus áreas de interés incluyen, Reconocimiento Lógico Combinatorio de Patrones, Inteligencia Artificial y Minería de datos y de textos. Cuenta con el reconocimiento de Perfil Deseable PRODEP de la SEP en México.



**María Esther Guevara Cruz**, recibió el título de Licenciada en Computación y el grado de Maestra en Ciencias de la Computación por la Benemérita Universidad Autónoma de Puebla, México. Ha participado en el desarrollo de proyectos de investigación en las áreas de la Inteligencia Artificial y el Reconocimiento de Patrones y sus aplicaciones.

Actualmente es Profesora de Tiempo Completo en la Universidad Tecnológica Fidel Velázquez, México. Sus áreas de interés incluyen, Reconocimiento Lógico Combinatorio de Patrones, Inteligencia Artificial y Minería de Datos y de Textos.



**Héctor Rafael Orozco Aguirre** es Profesor de Tiempo Completo del Centro Universitario UAEM Valle de México de la Universidad Autónoma del Estado de México, México. Obtuvo su Maestría en 2006 y Doctorado en 2010 en el Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Guadalajara, México. Como parte

de sus estancias de investigación en el extranjero, de 2007 a 2008 estuvo en el Laboratorio de Realidad Virtual de la Escuela Politécnica Federal de Lausana, Suiza, y de 2011 a 2012 en el Instituto de Innovación en Medios de Comunicación de la Universidad Tecnológica de Nanyang, Singapur. En 2011, fue galardonado por la Sociedad Mexicana de Inteligencia Artificial a la Tercera Mejor Tesis de Doctorado a nivel nacional. Actualmente, trabaja en Proyectos de Investigación referentes a Simulación de Estrategias de Predicción y Anticipación de Crimen para su Control y Reducción, Tutores Virtuales para la Mejora de la Enseñanza Educativa, así como Análisis, Modelado y Simulación de Tráfico Vehicular y Comportamiento Peatonal. Ha dirigido tesis de maestría y licenciatura, Ha sido conferencista magistral y ponente en diversos eventos y ha publicado artículos y capítulos de libro en las áreas de Inteligencia Artificial Aplicada e Inteligencia Computacional. Cuenta con el reconocimiento de Perfil Deseable PRODEP de la SEP en México.



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 2.5 México.

# **ANÁLISIS DEL DESEMPEÑO DE REDES NEURONALES PROFUNDAS PARA SEGMENTACIÓN SEMÁNTICA EN HARDWARE LIMITADO**

Performance Analysis of Semantic Segmentation Deep  
Neural Networks for Limited Hardware

Oscar Alejandro Soto-Orozco<sup>1</sup>  
Oscar.soto@ieee.org

Alma Delia Corral-Sáenz<sup>1</sup>  
[adcorral@itchihuahua.edu.mx](mailto:adcorral@itchihuahua.edu.mx)

Claudia Elizabeth Rojo-González<sup>1</sup>  
cerujo@itchihuahua.edu.mx

Juan Alberto Ramírez-Quintana<sup>1</sup>  
[jaramirez@itchihuahua.edu.mx](mailto:jaramirez@itchihuahua.edu.mx)

<sup>1</sup>División de Estudios de Posgrado e Investigación, Tecnológico Nacional de México/ I.T. Chihuahua, México.

1 **Resumen.** Segmentación semántica consiste en encontrar objetos previamente definidos en una imagen digital y se aplica en tecnologías como vehículos autónomos, interacción humano-maquina, realidad aumentada, robótica, etc. Los modelos más comunes para llevar a cabo esta forma de segmentación son las redes totalmente convolucionales, ya que reportan los mejores desempeños en la detección de objetos. Sin embargo, la mayor parte de estas redes tienen alto costo computacional y requieren de computadoras costosas, por lo que han surgido recientemente modelos basados en estas redes, pero con baja complejidad en cálculos para que las aplicaciones de segmentación semántica se puedan implementar desde sistemas embebidos. Por lo tanto, para contribuir con este esfuerzo, se presenta en este artículo un análisis detallado las redes Enet, Mobilenet v2, ERFNet y ESPNet v2, las cuales son redes populares en la literatura que se pueden correr desde un sistema embebido. Con base en los resultados, se concluye que los métodos que reemplazan la convolución regular por factorizaciones como la convolución separada en profundidad y convoluciones dilatadas con diversas ramas y el uso de otras estrategias como convoluciones saltadas e interpolaciones articuladas reducen el costo computacional comparando las métricas generadas por cada red como la huella de memoria, la precisión y el tiempo que tarda en segmentar una sola imagen.

2 **Palabras clave.** Aprendizaje profundo, Segmentación semántica, Redes neuronales convolucionales, Procesamiento de imágenes y video.

**Abstract.** Semantic segmentation consists in finding previously modified objects in a digital image and is applied in technologies such as autonomous vehicles, human-machine interaction, augmented reality, robotics, etc. The most common models to carry out this form of segmentation are totally convolutional networks, since they report the best performances in the detection of objects. However, most of these networks have high computational cost and are affected by expensive computers, so they have recently operated specific models in these networks, but with low complexity in analysis for semantic segmentation applications can be implemented from embedded systems. Therefore, to contribute to this effort, this article presents a detailed analysis of the Enet, Mobilenet v2, ERFNet and ESPNet v2 networks, which are popular networks in the literature that can run from an embedded system. Based on the results, it is concluded that the methods that replace the regular convolution by factorizations such as the deep convolution and dilated convolutions with various branches and the use of other strategies such as skipped convolutions and articulated interpolations reduce the computational cost by comparing the metrics generated for each red as the memory footprint, the precision and the time it takes to segment a single image.

**Keywords.** Deep learning, Semantic segmentation, Convolutional neural networks, Image and video processing

## 1. Introducción

El aprendizaje profundo con redes neuronales convolucionales (CNN por sus siglas en inglés) ha atraído la atención en los últimos años debido a que permite analizar y modelar características abstractas de forma automática, y genera mejores resultados que otros algoritmos supervisados y no supervisados en visión por computadora (Lockheed Martin, 2018).

Este tipo de redes son modelos de tipo *feedforward* (Shelhamer, Long, & Darrell, 2017) y tienen como entrada imágenes RGB de alta definición que se propagan hacia adelante a través de diversas neuronas con pesos y parámetros de aprendizaje. La arquitectura de estas redes se basa en módulos para extraer características y capas para clasificación. Los módulos de características se componen al menos una de las siguientes capas:

*Convolución.* Neuronas que se componen de pesos o filtros que se convolucionan con la salida de las capas más bajas para obtener información abstracta de la imagen.

*Agrupamiento.* Reducen la dimensión de la salida de una capa convolucional partiéndola en ventanas y seleccionando solo un valor que represente los elementos de dicha ventana.

*Normalización por lote.* Se normalizan los resultados de la capa anterior para evitar valores muy grandes o fuera de rango e incrementar la estabilidad de la red (Ioffe & Szegedy, 2015).

*Funciones de activación.* Definen la respuesta de una neurona y se modelan como sigmoideas, de base radial o de umbral.

Conforme la información de la entrada se propaga por la CNN hacia adelante, los módulos para extraer características encuentran información de colores, bordes y texturas cada vez más significativa y detallada. Las capas de clasificación se componen de neuronas similares a las de cualquier red neuronal *feedforward* y categorizan la información de la imagen de entrada para asignar una categoría semántica a cada uno de los píxeles de la imagen.

En la actualidad, existen una gran cantidad de arquitecturas de CNNs aplicadas a segmentación semántica (SS) que asignan una etiqueta  $l$  de un conjunto comprendido por  $\mathcal{L} = \{l_1, l_2, \dots, l_n\}$  para cada elemento del conjunto de variables aleatorias de entrada  $X = \{x_1, x_2, \dots, x_N\}$ . Cada etiqueta  $l$  representa una clase u objeto con significado semántico diferente como avión, vehículo, señal de tráfico o fondo. Por otro lado, la variable  $X$  es una imagen de una secuencia de video. A menudo se etiquetan las clases con máscaras de color para visualizarlas en las imágenes, como se puede observar en la Figura 1.



Figura 1. Ejemplos de segmentación semántica con base de datos Cityscapes (Cordts et al., 2016)

Existen arquitecturas profundas de CNNs populares en el estado del arte debido al desempeño que obtuvieron en la base de datos de ImageNet Large Scale Visual Recognition Challenge (ILSVRC sus siglas en inglés), la cual es utilizada formalmente en la literatura para realizar pruebas, ya que tiene más de 190,000 imágenes con sus *groundtruths* para clasificar diversos objetos (Krizhevsky, Hinton, & Sutskenvet, 2012).

La primera de estas arquitecturas fue AlexNet (Krizhevsky et al., 2012), una red de 8 capas ocultas compuestas por 5 convolucionales y 3 totalmente conectadas que logró un error de precisión de 37.5% en el conjunto de prueba de la base de datos de ImageNet. Este error fue considerablemente menor que el reportado en los métodos del estado del arte en ese momento.

Otra red popular es la Visual Geometry Group (VGG) (Simonyan & Zisserman, 2014), propuesta por el departamento de ingeniería y ciencia de la universidad de Oxford y se conforma de 16 capas convolucionales y 3 capas totalmente conectadas. La principal diferencia de VGG sobre AlexNet es que, en vez de tener pocas capas con grandes campos receptivos, utiliza una pila de capas de convolución con pequeños campos receptivos en las primeras capas.

Este uso en las capas de convolución tuvo tal impacto, que en los años posteriores sería la base para GoogleNet, una red propuesta en (Szegedy et al., 2015) por investigadores de la universidad de Carolina del norte y la universidad de Michigan en colaboración con Google Inc. Esta red tiene un total de 22 capas compuestas por convoluciones con agrupamientos y capas totalmente conectada, y agrega un mecanismo de aprendizaje basado en una regla hebbiana.

Además, tiene un procesamiento de múltiples escalas para generar la capacidad de adquirir información en contexto del escenario. GoogleNet es muy profunda y computacionalmente costosa, ya que es necesario desarrollar una gran cantidad de capas y filtros que conllevan muchas operaciones de punto flotante por segundo para obtener resultados aceptables en la literatura científica. Este número de capas exige una enorme cantidad de recursos como núcleos especializados, memoria dedicada a video y altas velocidades de procesamiento.

Estas redes se han utilizado para proponer modelos de SS funcionando con buenos desempeños y de forma eficiente en computadoras como estaciones de trabajo o computadoras de alta capacidad de procesamiento.

Esto causa que el desarrollo tecnológico en las áreas como vehículos autónomos, interacción humano-maquina, realidad aumentada y robótica sea sumamente costosa (Lateef & Ruichek, 2019).

Por ello, en los últimos años se han generado líneas de investigación en CNNs que se centran en analizar arquitecturas de redes para optimizar el uso de los recursos de procesamiento y permitir la implementación de SS en sistemas embebidos o plataformas de hardware a baja potencia.

A partir de esta línea, se han generado modelos de SS que funcionan de forma exitosa en plataformas embebidas de hardware limitado y han permitido ahorrar costos de forma significativa en el desarrollo de tecnología. Por lo tanto, para contribuir en este esfuerzo de desarrollar esquemas de SS en hardware limitado, en este artículo se propone un análisis detallado del desempeño de los modelos más populares en la literatura para SS que tienen un número reducido de operaciones de punto flotante (FLOPS), bajo uso de memoria y bajo consumo de energía.

El objetivo de este análisis es estudiar las estrategias de diseño que permiten el desarrollo de redes eficientes, y proveer un punto de partida para desarrollar arquitecturas novedosas cada vez más eficientes y que permitan desarrollar tecnología a costos significativamente bajo.

El resto del artículo presenta de la siguiente forma: la sección 2 describe diversas arquitecturas de CNNs para SS que son eficientes en el uso de recursos computacionales. La sección 3 analiza los diversos parámetros que ayudan a colocar a estas arquitecturas como modelos eficientes en recursos. La sección 4 presenta los resultados del análisis realizado a estos modelos y la sección 5 presenta las conclusiones.

## 2. Redes totalmente convolucionales para segmentación semántica

El éxito en desempeño que han tenido las CNNs en SS se debe a que estas encuentran características abstractas que definen a cada clase semántica. Por esta razón, en la actualidad se han generado múltiples arquitecturas de CNNs para de SS, entre las cuales, las más populares son las redes totalmente convolucionales (FCN por sus siglas en inglés)(Sevak, Kapadia, Chavda, Karungan, & Sujatha, 2017).

La primera FCN fue propuesta por Jonathran Long y Evan Shelhamer de la Universidad de Berkeley (Shelhamer, Long, & Darrell, 2017), y se basa en tomar ventaja de las arquitecturas existentes de CNNs de clasificación de imágenes para aprender características de diferente jerarquía. Para ello, la FCN tiene cinco bloques, como se ve en la Figura 2, donde el primer bloque es la imagen de entrada. El segundo bloque es el codificador y se compone de un conjunto de capas convolucionales que comprimen la información de la imagen de entrada  $X$  para extraer características.

El tercer bloque es una imagen comprimida  $F(X)$  que contiene el mapa de características y es la entrada para el cuarto bloque, el cual se le conoce como decodificador. Este bloque se basa en un conjunto de capas de desconvolución que descomprimen los mapas de características para obtener una imagen que contiene la segmentación de los objetos contenidos en la imagen de entrada  $X'$ . Las capas de este último bloque se componen de desconvoluciones para obtener como salida mapas de características descomprimidos. La última capa del decodificador obtiene la combinación de todos los mapas de características generados por las capas anteriores para ser interpolados con la convolución transpuesta descrita por (Taylor, 2010), que toma la imagen de salida de la capa anterior  $Y^{i-1}$ , compuesto por  $k_0$  canales  $Y_1^i, \dots, Y_{k_0}^i$ , lo siguiente:

$$\sum_{k=1}^{k_1} Y_k^{i-1} \oplus f_{k,c} = y_c^i \quad (1)$$

Donde  $f$  son los filtros, y la capa genera una salida  $Y_c^i$  donde  $c = 1, \dots, k_1$ . Es decir, la FCN es una CNN que tienen un bloque de decodificación en vez de capas de clasificación. De esta manera se produce una imagen reconstruida  $X'=(G \circ F)(X)$  que tiene la misma resolución de la imagen de entrada y contiene los resultados de la segmentación, donde cada elemento de  $X'$  tiene asignada una categoría semántica que indica a que clase pertenece cada pixel de la entrada.

Al modelo de la FCN también se le conoce como red autocodificadas.



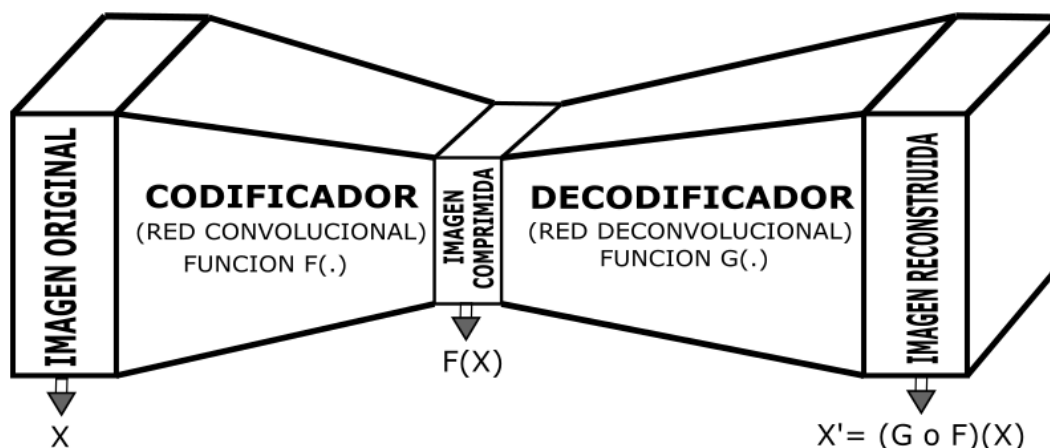


Figura 2. Estructura básica de red convolucional auto codificada.

En otras palabras, la FCN toma la imagen original y la comprime en mapas de características que contienen la información semántica de los objetos, pero con resolución espacial y de color reducidas.

Conforme la red se torna más profunda, se comprimen más los mapas de características, y cuando se propaga estos mapas por todo el codificador, se descomprimen en resolución espacial en el decodificador hasta llegar al tamaño de  $X$ . (Neapolitan & Neapolitan, 2018).

Esta FCN mostró un mejor desempeño para SS que los métodos estadísticos de aprendizaje automático ya que las capas convolucionales extraen características abstractas al igual que las capas de las CNNs mencionadas en la introducción y las capas de deconvolucion densifican las activaciones obtenidas por el codificador generando una activación ampliada, pero con una mayor densidad de píxeles.

### 3. Evaluación de modelos eficientes

A pesar de la potencia, buena precisión y flexibilidad del modelo FCN, este tiene diversos aspectos que dificultan su aplicación ya que no considera la variación espacial que pueda existir en los objetos, ni la relación que mantienen las diferentes clases en el escenario y no siempre se puede ejecutar en tiempo real con imágenes de alta resolución en sistemas embebidos.

A pesar de ello, en la literatura se reportan cuatro modelos para SS inspirados en la FCN, cuya arquitectura en el codificador y el decodificador les permiten desempeños aceptables en precisión y uso de recursos.

Este desempeño se puede describir en las siguientes métricas:

**Tiempo de inferencia:** es el total de operaciones de punto flotante por segundo (FLOPS) necesarias para ejecutar el modelo. Los FLOPS son proporcionales al tiempo en que se propaga la entrada por toda la red y a las especificaciones del hardware en el que se trabaje. Para el caso de este análisis, el tiempo de inferencia se obtuvo mediante pruebas realizadas en una laptop con procesador Intel Core i7 de 8va generación con una GPU GTX 1060 con 6 Gb de memoria dedicada a video.

**Huella de memoria:** se refiere a la cantidad de memoria requerida para procesar la FCN y obtener la máscara de SS.

**Precisión:** consiste en el porcentaje de similitud que tiene el resultado de segmentación obtenido por el modelo con un objetivo de segmentación o *groundtruth*. Generalmente, esta métrica es la más utilizada en las bases de datos de SS y se define por la intersección sobre la unión IoU.

Los modelos seleccionados poseen un desempeño equilibrado en estas 3 métricas ya que mantienen una precisión apropiada, una huella de memoria baja y un tiempo de ejecución en tiempo real en sistemas que utilicen procesadores ARM embebidos o computadoras con fuentes de poder de menos de 100 W.

Este tipo de hardware limitado es de bajos recursos computacionales y de acuerdo con un estudio de tecnología, puede generar costos significativamente más bajos en el desarrollo de aplicaciones de SS. De esta manera, fueron seleccionados para este análisis los modelos de Enet (Paszke, Chaurasia, Kim, & Culurciello, 2016), Mobilenet v2 (Sandler *et al.*, 2017), ESPNet v2 (Mehta, Rastegari, Caspi, Shapiro, & Hajishirzi, 2018), ERFNet (Romera, Álvarez, Bergasa, & Arroyo, 2018) y FastFCN (H. Wu, 2019). A continuación, se describen cada una de estas redes.

En (Paszke *et al.*, 2016), se propone una FCN denominada Red eficiente (Enet por sus siglas en inglés), la cual tiene una arquitectura autocodificada de 23 capas. El módulo codificador se compone de 20 capas de conexiones residuales y convoluciones factorizadas que aumentan la eficiencia, mientras se mantiene un desempeño aceptable.

El módulo decodificador se compone de 3 capas deconvolucionales que aumentan la resolución de los mapas de características obtenidos por la sección codificadora. ENet logró una velocidad de procesamiento de 83 cuadros por segundo en una GPU NVIDIA Titan y 7 cuadros por segundo en una JetsonTX1.

Este modelo fue retomado por (Sandler *et al.*, 2018), y se propuso una versión llamada Mobilenet v2, de arquitectura codificador-decodificador con módulos de pirámide de convoluciones dilatadas (ASPP), en donde se obtiene información de contexto (Mehta, Rastegari, Shapiro, & Hajishirzi, 2018) a diferentes niveles de dilatación.

También se utilizan convoluciones de 1x1 para agrupar características de la imagen y un *stride* o paso de salida que controle la resolución del mapa de segmentación. De esta manera, la extracción de características que proporciona Mobilenet v2 propicia un tiempo de inferencia y huella de memoria relativamente bajos en comparación a los de otros modelos, mientras la precisión del sistema permanece aceptable.

Esto es gracias a que la red está compuesta principalmente por módulos de convolución separables basadas en cuellos de botella, el cual consiste en generar una convolución 1x1 para aumentar la dimensión del canal, luego se hace una convolución separada en profundidad, realizando un filtrado ligero mediante la aplicación de un único filtro convolucional por canal de entrada para reemplazar las convoluciones regulares, en esencia ambas convoluciones, la regular y la separada por profundidad funcionan de la misma manera pero reduciendo significativamente la complejidad computacional.

Finalmente, se realiza otra convolución de 1x1 para reducir la dimensión del canal construyendo nuevas características al calcular combinaciones lineales de los canales de entrada. En la Figura 3. se aprecia este modelo de cuello de botella. La arquitectura de Mobilenet v2 contiene inicialmente una capa convolucional con 32 filtros, seguidos de 19 capas de cuello de botella previamente descritos.

Para evitar la no linealidad en la salida de los filtros, se usó la rectificación lineal ReLU6 (He et al., 2018). Con base en el estándar de las redes actuales, Mobilenet utiliza filtros de 3x3, un *dropout* (operación donde las neuronas generan un aprendizaje parcial de la red para evitar el sobre entrenamiento) y normalización de lote durante el entrenamiento.

La experimentación que se tuvo con este modelo mostró que para redes menos profundas genera resultados más precisos, mientras que las redes de mayor profundidad tienen un rendimiento ligeramente más preciso al segmentar con filtros más profundos.

En (Romera et al., 2018), se propone la red factorizada residual eficiente (ERFNet por sus siglas en inglés). El elemento central de esta arquitectura es el diseño de una capa novedosa que aprovecha convoluciones con los núcleos de una dimensión y las conexiones saltadas mostradas en la Figura 4. donde la entrada se suma con el resultado de las capas convolucionales.

Mientras que las conexiones de salto permiten que las convoluciones aprendan funciones residuales que facilitan el entrenamiento, las convoluciones factorizadas permiten una reducción significativa de los costos computacionales, pero manteniendo una precisión similar a las redes con convoluciones de dos dimensiones (Asadi, Chen, Han, Wu, & Lobaton, 2019).

El bloque propuesto se apila secuencialmente para construir la arquitectura de la forma autocodificador, que produce la segmentación semántica de extremo a extremo, como se muestra en la Figura 5. en una visualización de la arquitectura de ERFNet.

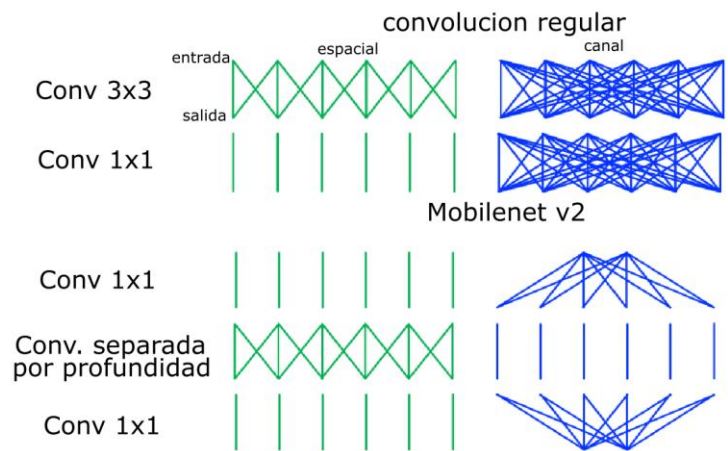


Figura 3. Diferencias entre las convoluciones de 1x1 y 3x3 regulares y el cambio de canal en arquitectura Mobilenet v2 de forma de cuello de botella

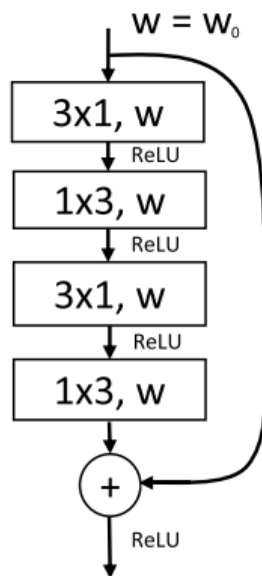


Figura 4. Representación de capas residuales,  $w$  representa el número de mapas de características que entran a cada capa, reducidos internamente en 4 en el diseño de cuello de botella.

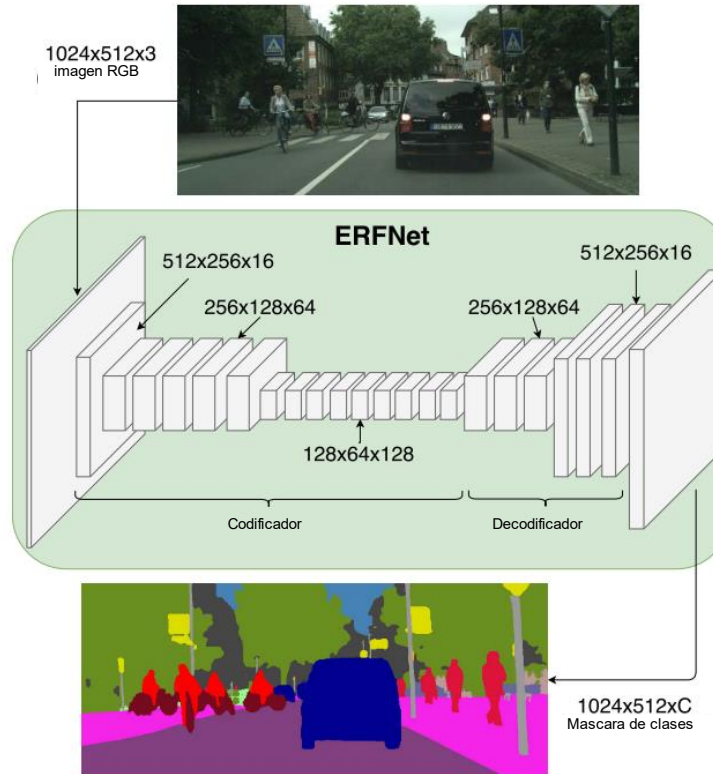


Figura 5. Arquitectura de ERFNet, con una imagen de alta resolución como entrada generando una máscara de segmentación de la misma resolución (Romera et al., 2018).

Por otro lado, en (Mehta, Rastegari, Shapiro, et al., 2018) se presentó una segunda versión de la red de pirámide espacial eficiente de convoluciones dilatadas (ESPNet por sus siglas en inglés), una red eficiente en términos de procesamiento, memoria y potencia. Esta arquitectura se muestra en la Figura 6. el bloque principal de construcción es un bloque convolucional nuevo llamado pirámide espacial eficiente (ESP por sus siglas en inglés) y se basa en descomponer una convolución estándar en dos pasos:

1. *Convolución puntual.* Consiste en el uso de un kernel de 1x1 para proyectar mapas de características de alta dimensión en un espacio de baja dimensión.
2. *Pirámide de convolución dilatada espacial.* Es un mecanismo similar al de ASPP que aprovecha las convoluciones atrous mostradas en la Figura 7. Pero en este caso, es menos costoso computacionalmente debido al acomodo en forma de pirámide, ya que el módulo ASPP utiliza una serie de convoluciones dilatadas sin orden jerárquico y tiene la capacidad de usar diferentes niveles espaciales para aprender representaciones.

Esta ayuda a decodificar los mapas de características descomprimiéndolos para aprender diferentes representaciones de los campos semánticos a grandes campos receptivos (cantidad de información visual que puede adquirir cada neurona o kernel) efectivos, usando  $N$  kernels de una dimensión  $n \times n$  de convoluciones dilatadas de forma simultánea, cada uno con una relación de  $2^{N-1}, N = \{1, \dots, N\}$  que genera el total de parámetros  $p$ .

Esta factorización reduce drásticamente el número de parámetros y memoria requeridos por el módulo, mientras se mantiene un gran campo receptivo, teniendo un total de parámetros:

$$p = [(n - 1)2^{N-1} + 1]^2 \quad \begin{matrix} (2 \\ ) \end{matrix}$$

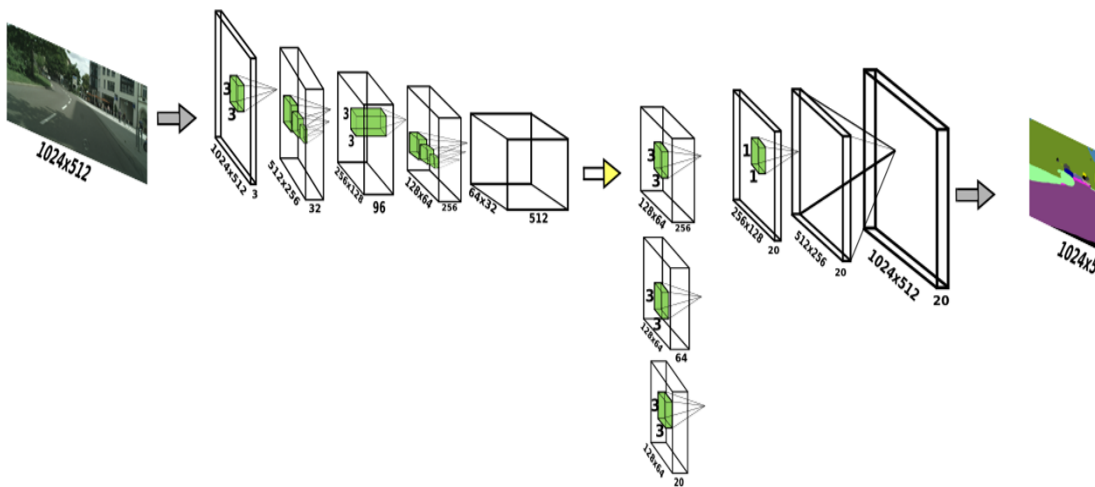


Figura 6. Estructura general de arquitectura ESPNet V2

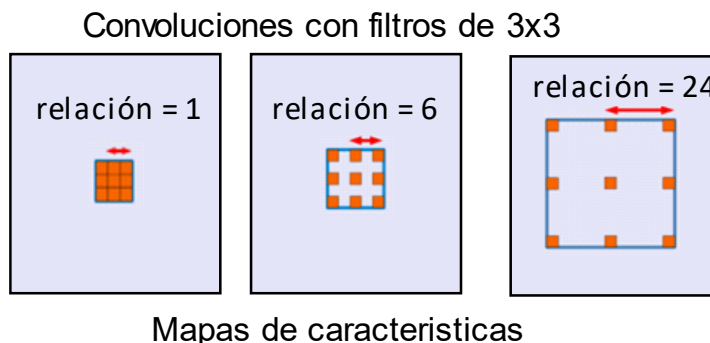


Figura 7. Convoluciones atrous con kernel de 3x3 de diferentes relaciones, la convolución regular corresponde al recuadro de la izquierda (Chen, Papandreou, Schroff, & Hartwig, 2017).

Esta operación de pirámide convolucional es llamada pirámide espacial de convoluciones dilatadas, porque cada kernel aprende pesos con diferentes campos receptivos y se asemeja a una pirámide espacial. Además, estos módulos ESP tienen un divisor de ancho  $N$  (total de filtros utilizados por capa) para reducir el costo computacional. Este parámetro tiene como propósito:

1. Reducir la dimensionalidad de las capas de características de manera uniforme en cada módulo ESP de la red. Es decir, por cada  $N$  los módulos reducen de un espacio dimensional  $M$  a un espacio  $\frac{R}{K}$  usando la convolución puntual.
2. Dividir los espacios de características de baja dimensión en  $K$  ramas paralelas.
3. Transformar cada rama procesa los mapas de características de forma simultánea usando *kernels* de convoluciones dilatadas de  $n \times n$  a diferentes relaciones de dilatación definidas por:

$$2^{N-1}, N = \{1, \dots, N - 1\} \quad (3)$$

4. Unir la salida de los kernel convolucionales dilatados paralelos  $K$  se concatena para producir un mapa de características de salida de dimensión  $N$ .

Finalmente, (Wu, Zhang, Huang, Liang, & Yu, 2019) presentaron la red totalmente convolucional rápida (FastFCN por sus siglas en inglés), y se trata de una arquitectura para SS que parte del modelo FCN la cual se muestra en la Figura 8. Utiliza convoluciones dilatadas en su estructura principal para extraer mapas de características de alta resolución. Para evitar una complejidad computacional pesada y una huella de memoria grande, se propone un módulo para descomprimir los mapas de características llamado pirámide de muestreo ascendente conjunta (*JPU por sus siglas en ingles*). Este módulo tiene como objetivo para una imagen de entrada baja resolución y un *groundtruth* de alta resolución generar una salida de alta resolución mediante la transferencia de detalles y estructuras de la imagen de entrada. Generalmente, el *groundtruth* de baja resolución  $Z_1$  es generada empleando la transformación  $F(X)$  en la imagen de baja resolución de entrada  $X_1$ , por ejemplo,  $Z_1 = f(X_1)$ . Teniendo  $X_1$  y  $Z_1$  es necesario obtener la transformación  $\hat{F}(X)$  para aproximar  $F(X)$ , donde la complejidad computacional de  $\hat{F}(X)$  es menor que  $F(X)$ . Por ejemplo, si  $F(X)$  es un perceptrón multi capa (MLP), entonces  $\hat{F}(\cdot)$  puede ser simplificada como una transformación lineal. La imagen *grountruth* de alta resolución  $y_h$  es obtenida aplicando la transformación  $\hat{F}(\cdot)$  en la imagen de alta resolución de entrada  $X_h$  de la forma  $Z_h = \hat{F}(X_h)$  formalmente dadas  $X_1$ ,  $Z_1$  y  $X_h$  el muestreo ascendente conjunto es definido como:

$$Z_h = \hat{f}(x_h), \text{ Donde } \hat{f}(\cdot) = \min_{h(\cdot) \in \varphi} ||y_1 - h(x_1)|| \quad (4)$$

Donde  $\varphi$  es el conjunto de todas las funciones de transformación posibles, y  $||\cdot||$  es una métrica de distancia. El resto de la pirámide está formada por convoluciones y convoluciones dilatadas como se muestra en la Figura 9.

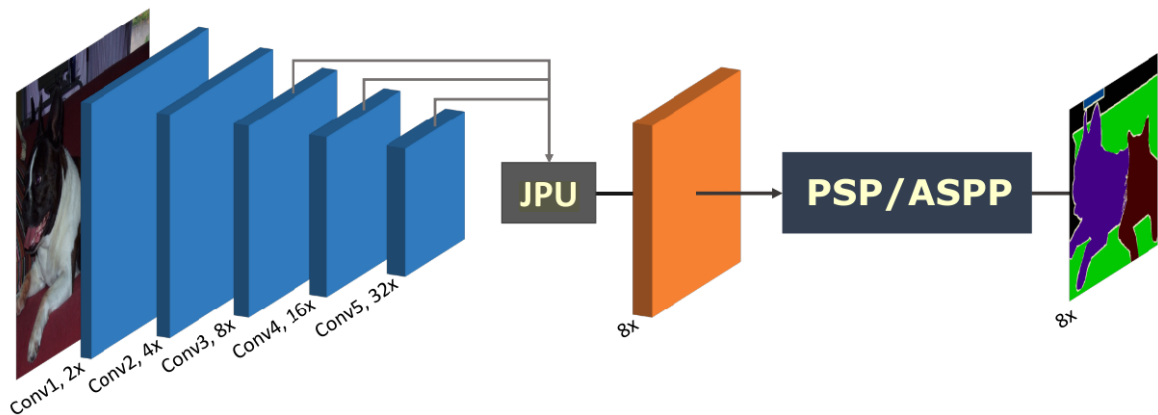


Figura 8. La arquitectura FastFCN que utiliza como base el modelo FCN, con un modelo novedoso llamado pirámide de interpolación articulada (JPU) que ayuda a interpolar y combinar las características obtenidas (Wu et al., 2019).

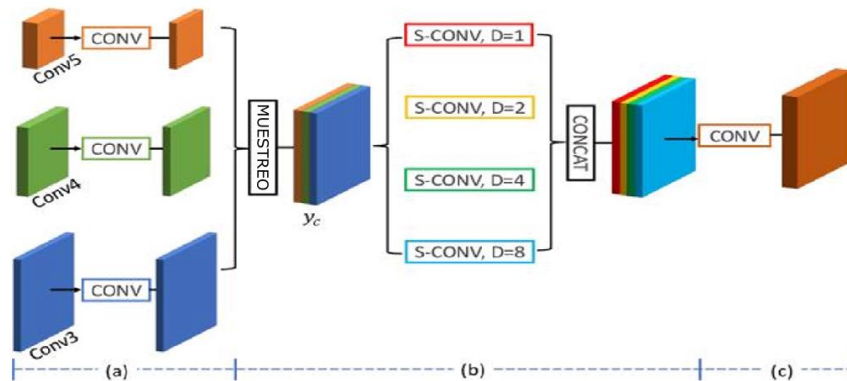


Figura 9. Modulo JPU visto de mejor perspectiva a colores.

#### 4. Resultados

Las FCNs mencionados reportan buenos desempeños en la literatura con diversas base de datos como CamVid (Amoros et al., 2008), Mapillary Vistas (Neuhold, Ollmann, Bulo, & Kotschieder, 2017), Cityscapes (Cordts et al., 2016) y ADE20K (Zhou et al., 2017). CamVid fue la primera base de datos oficial para estudiar la SS y se compone de 700 imágenes grabadas desde un vehículo y tiene etiquetas clasificadas en 32 clases que representan objetos de un entorno urbano como edificios, caminos, vegetación, animales, luces de tráfico, ciclistas, trenes etc.

Mapillary Vistas tiene un total de 25,000 imágenes de alta resolución de diversas partes del mundo con etiquetas específicas para 37 clases. Cityscapes tiene un total de 5,000 *groundtruth* detallados de 50 ciudades diferentes y se generaron 30 clases diferentes divididas en 8 categorías que representan objetos urbanos. ADE20K es una de las bases de datos más extensa, ya que tiene un total de 20,000 imágenes de entrenamiento y 2,000 para validación. Cuenta con un total de 150 objetos etiquetados donde destacan entornos urbanos y diversos escenarios como hogares.



Entre los objetos contenidos en esta base de datos se encuentran camas, cuadros de pintura, muebles, artículos del hogar, etc.

Como se mencionó previamente, cada modelo fue implementado en diferentes plataformas de hardware y con diferentes bases de datos, por lo que no se puede hacer una evaluación objetiva de los modelos de su funcionamiento.

Por lo tanto, para realizar una evaluación del desempeño, en esta sección se presentan los resultados de un análisis de la precisión, tiempo de ejecución y huella de memoria de Enet, Mobilenet v2, ESPNet v2, ERFNet y FastFCN.

Esta evaluación se realizó con la base de datos de Cityscapes (Cordts et al., 2016), ya que es la más popular en la literatura y tiene un *benchmark* (prueba de rendimiento o comparativa de un sistema) donde compiten diferentes modelos entrenados y probados con las diferentes imágenes de esta base de datos. La métrica utilizada es la PASCAL VOC *intersection-over-union* que consiste en obtener la relación:

$$IoU = \frac{TP}{(TP + FP + FN)} \quad (4)$$

donde TP, FP y FN son la cantidad de píxeles de verdadero positivo, falso positivo, y falso negativo respectivamente. La Tabla 1. muestra los resultados y comparación en tiempo de inferencia, promedio de precisión y huella de memoria en la base de datos Cityscapes. Respecto a los campos de la Tabla 1.

Precisión se tomó con la métrica IoU, el tiempo de inferencia se midió con la librería *time* de python y la huella de memoria se midió al propagar por la red una imagen de 3 canales a una resolución de 1024 por 512 píxeles, utilizando la herramienta *summary* de pytorch. En este caso, únicamente FastFCN (Wu et al., 2019) fue probado en ADE20K y se obtuvo una precisión de 55.84% con un tiempo de inferencia de .025 segundos. FastFCN fue probado en esta base para otorgar otro punto de vista a estos modelos, ya que esta base de datos en específico no se centra en ambientes urbanos lo que aporta un mayor panorama del potencial que este tipo de redes eficientes tiene para otro tipo de aplicaciones.

Tabla 1. Comparación de modelos seleccionados en precisión y velocidad de inferencia con la base de datos Cityscapes con procesador Intel core i7 8750H con una GPU Gtx 1060 de 6Gb de memoria para video.

	Precisión (IoU)	Tiempo de inferencia	Huella de memoria
Enet	0.513	0.05 s	6779.34 MB
Mobilenet v2	0.707	0.1 s	1616.50 MB
ERFNet	0.727	0.02 s	1917.87 MB
ESPNet v2	0.626	0.0089 s	2471.59 MB

El método con mayor precisión es ERFNet, ya que tiene un IoU más alto que el resto de los métodos. Esta precisión se debe a la profundidad que presenta en su codificador, ya que al generar un mayor número de mapas de características cada vez más profundas y toma más información contextual del escenario. Por otro lado, el método con menor precisión es Enet, y se debe principalmente a que la contribución de este se enfocaba en proponer las convoluciones dilatadas para reducir la complejidad computacional del modelo. Luego de Enet, se propusieron modelos con este tipo de convolución que mejoraban la precisión.

En cuanto tiempo de inferencia, ESPnet v2 tiene el menor tiempo de ejecución debido a la descomposición que presenta en convoluciones grupales, puntuales y dilatadas acomodadas en un esquema piramidal. Mobilenet v2 presenta el mayor tiempo en inferencia, lo cual se atribuye a la profundidad que presenta su arquitectura, además de no utilizar la operación de convolución dilatada.

En cuanto a huella de memoria, el modelo que presenta la memoria total más baja fue Mobilenet v2 gracias a la poca profundidad que mantiene el modelo. Por otro lado, el modelo con más huella de memoria es Enet debido principalmente a que utiliza un mayor número de convoluciones regulares con filtros de más de 3 pixeles de dimensión.

Un análisis cualitativo de los modelos seleccionados se muestra en la Figura 10. donde se ven reflejados las puntuaciones de precisión alcanzados por cada modelo, de izquierda a derecha se puede observar la imagen de entrada seguida de la imagen objetivo o *groundtruth*, posteriormente los resultados de Enet, ERFNet, ESPNet v2 y MobileNet v2 donde es posible destacar en cada imagen la forma en que se segmentan objetos distintos de interés, en la primera imagen observamos como MobileNet v2 tiene un resultado mas limpio aunque omite el fondo negro ubicado en la esquina superior izquierda que se encuentra en el *groundtruth*, por otro lado en la segunda imagen es ERFNet el que logra segmentar con una mayor

precisión el objeto que le cuesta mas al resto de modelos, que en este caso es el camión blanco que pasa por delante de la camara.

Por ultimo la tercera imagen muestra un paso peatonal donde encontramos distintos objetos destacando las personas y los vehículos del fondo para este caso ERFNet y ESPNet v2 presentan los resultados mas parecidos al *groundtruth* de esta imagen. Estos resultados muestran una correlación directa con las puntuaciones generales mostradas anteriormente, y se logra observar los objetos que resultan ser mas complicados para segmentar por cada modelo presentado en este análisis.

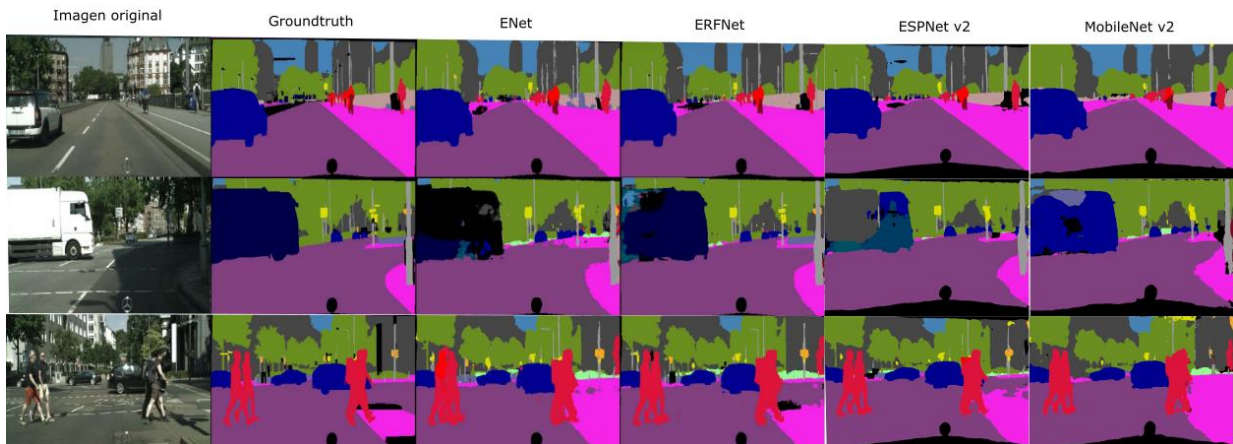


Figura 10. Análisis cualitativo de modelos comparando la imagen original, su *groundtruth* y cada una de las máscaras de segmentación generadas por cada modelo.

## 5. Conclusiones

Este artículo presenta un análisis del desempeño de las redes FCN eficientes enfocadas a segmentación semántica para hardware limitado. Las redes que fueron seleccionadas son Enet, Mobilnet v2, ERFNet, ESPNet v2 y FastFCN y las métricas utilizadas para el análisis son precisión, tiempo de inferencia y huella de memoria.

Estos modelos fueron evaluados con la base de datos Cityscapes y los resultados mostraron que los mejores métodos para SS fueron ERFNet gracias a su alta precisión obtenida validando el método, ESPNet v2 debido al poco tiempo que requiere para inferir el resultado de segmentación por cada imagen de entrada y Mobilenet v2 gracias a su reducida huella de memoria generado por los parámetros de la red.

Los buenos resultados en precisión de ERFNet se deben a que su modelo de baja profundidad y convoluciones con salto residual permiten una rápida convergencia en el entrenamiento, lo cual implica que la red siempre reduce la función de costo hasta encontrar los valores de los pesos de las neuronas que generan el error mínimo. Otro aspecto que ayuda a ERFNet es que el módulo JPU descomprime con una menor cantidad de FLOPS que las demás redes y logra generalizar los mapas de características obtenidos por las capas posteriores.

Por otro lado, ESPNet, generó una baja cantidad de FLOPS debido a la baja complejidad computacional obtenida por la estructura tipo ASPP. No obstante, este modelo no tuvo la menor huella de memoria ya que genera una gran cantidad de información con las convoluciones dilatadas en forma de pirámide.

De esta manera, se puede ver que diversas propuestas para la arquitectura de las redes influyen en la precisión o en el tiempo de inferencia. A futuro, los modelos de FCN eficientes se deben enfocar en mantener un equilibrio en el total de filtros utilizados en cada capa, haciendo una inspección del total de parámetros generados con herramientas que permitan el ambiente en el que se esté desarrollando el modelo.

Una estrategia clara presente en los trabajos analizados fue el evitar usar convoluciones completas ya que el total de parámetros generados por esta operación incrementa de forma exponencial, mientras que las ASPP y las convoluciones factorizadas grupales y puntuales reducen significativamente el total de parámetros generados.

El único problema con estas técnicas es la pérdida de información generada por los huecos y los puntos de la imagen en donde no se convoluciona el filtro. Por lo tanto, es necesario incrementar y reestructurar los filtros dilatados de manera que se obtenga la información de contexto de cada objeto en su totalidad.

## Referencias.

- Amoros, P., Balsells, M. A., Buisan, M., Byrne, S., Fuentes-Pelaez, N., & Gabriel J. Brostow a, b,\*, Julien Fauqueur a, R. C. a. (2008). Semantic object classes in video: A high-definition ground truth database. *Revista de Cercetare Si Interventie Sociala*, 42(2), 120–144.  
<https://doi.org/10.1016/j.patrec.2008.04.005>
- Asadi, K., Chen, P., Han, K., Wu, T., & Lobaton, E. (2019). *Real-time Scene Segmentation Using a Light Deep Neural Network Architecture for Autonomous Robot Navigation on Construction Sites*. Retrieved from <http://arxiv.org/abs/1901.08630>
- Chen, L., Papandreou, G., Schroff, F., & Hartwig, A. (2017). *Rethinking Atrous Convolution for Semantic Image Segmentation*.
- Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., ... Schiele, B. (2016). *The Cityscapes Dataset for Semantic Urban Scene Understanding*. <https://doi.org/10.1109/CVPR.2016.350>
- He, K., Zhang, X., Yang, H., Han, K., Zhu, D., Lun, P., & Zhao, Y. (2018). Delving Deep into Rectifiers: Surpassing Human-Level performance on imagenet classification. *Biochemical and Biophysical Research Communications*, 498(1), 254–261. <https://doi.org/10.1016/j.bbrc.2018.01.076>
- Ioffe, S., & Szegedy, C. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Lateef, F., & Ruichek, Y. (2019). Survey on semantic segmentation using deep learning techniques. *Neurocomputing*, 338, 321–348.  
<https://doi.org/10.1016/j.neucom.2019.02.003>
- Mehta, S., Rastegari, M., Caspi, A., Shapiro, L., & Hajishirzi, H. (2018). ESPNet: Efficient spatial pyramid of dilated convolutions for semantic segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11214 LNCS, 561–580. [https://doi.org/10.1007/978-3-030-01249-6\\_34](https://doi.org/10.1007/978-3-030-01249-6_34)
- Mehta, S., Rastegari, M., Shapiro, L., & Hajishirzi, H. (2018). *ESPNetv2: A Light-weight, Power Efficient, and General Purpose Convolutional Neural Network*. Retrieved from <http://arxiv.org/abs/1811.11431>
- Neapolitan, R. E., & Neapolitan, R. E. (2018). Neural Networks and Deep Learning. In *Artificial Intelligence*. <https://doi.org/10.1201/b22400-15>
- Neuhold, G., Ollmann, T., Bulo, S. R., & Kotschieder, P. (2017). The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes. *Proceedings of the IEEE International Conference on Computer Vision, 2017-Octob*, 5000–5009. <https://doi.org/10.1109/ICCV.2017.534>
- Paszke, A., Chaurasia, A., Kim, S., & Culurciello, E. (2016). *ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation*. 1–10.
- Romera, E., Álvarez, J. M., Bergasa, L. M., & Arroyo, R. (2018). *ERFNet: Efficient Residual Factorized ConvNet for Real-Time Semantic Segmentation*. 19(1), 263–272.
- Sandler, M., Zhu, M., Zhmoginov, A., & Apr, C. V. (2018). *MobileNetV2: Inverted Residuals and Linear Bottlenecks*.

- Sevak, J. S., Kapadia, A. D., Chavda, J. B., Karungan, D., & Sujatha, N. (2017). Survey on Semantic Image Segmentation Techniques. *Proceedings of the International Conference on Intelligent Sustainable Systems, 4(Iciss)*, 306–313. Retrieved from [www.jetir.org](http://www.jetir.org)
- Shelhamer, E., Long, J., & Darrell, T. (2017). *Fully Convolutional Networks for Semantic Segmentation*. 39(4), 640–651.
- Taylor, G. W. (2010). *Deconvolutional Networks slides*. 2528–2535. <https://doi.org/10.1109/CVPR.2010.5539957>
- Wu, H., Zhang, J., Huang, K., Liang, K., & Yu, Y. (2019). *FastFCN: Rethinking Dilated Convolution in the Backbone for Semantic Segmentation*. (1). Retrieved from <http://arxiv.org/abs/1903.11816>
- Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralla, A. (2017). Scene parsing through ADE20K dataset. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 5122–5130. <https://doi.org/10.1109/CVPR.2017.544>

## Notas biográficas



**Oscar Alejandro Soto Orozco.** Obtuvo el grado de Ingeniero en Electrónica del Instituto Tecnológico Nacional campus Chihuahua 2018 y actualmente se encuentra estudiando para obtener el grado de Maestro en Ciencias en Ingeniería Electrónica del Instituto Tecnológico de Chihuahua, su investigación es en el área de procesamiento digital de señales e imágenes, enfocado a segmentación semántica en video con hardware limitado.



**Alma Delia Corral Sáenz.** Recibió el título de Ingeniera en Sistemas Computacionales en Hardware de la Universidad Autónoma de Chihuahua en 1999 y el de Maestra en Ciencias en Ingeniería Electrónica del Instituto Tecnológico de Chihuahua en 2003. Actualmente es profesora y coordinadora del Doctorado y la Maestría en Ciencias en Ingeniería Electrónica en el mismo Instituto, y participa en trabajos de investigación de las áreas de procesamiento de señales y visión por computadora.



**Claudia Elizabeth Rojo-González.** Recibió el título de Contador Público en 1995 y la Maestría en Administración en 2000, ambos títulos en la Universidad Autónoma de Chihuahua. Actualmente es profesora y coordinadora de la Maestría en Administración de Negocios en el Instituto Tecnológico de Chihuahua, y participa en trabajos de investigación auxiliando a profesores de ingeniería en el desarrollando de modelos financieros para desarrollo tecnológico.



**Juan Alberto Ramírez Quintana.** Recibió los grados de ingeniería (2004), maestría (2007) y doctorado (2014) en ingeniería electrónica del Instituto Tecnológico de Chihuahua, México. Actualmente trabaja como profesor-investigador en el Instituto Tecnológico de Chihuahua, cuenta con diversas publicaciones en revistas y congresos y dirige varias tesis a nivel licenciatura maestría y doctorado. Sus áreas de interés son visión por computadora, procesamiento de señales, aprendizaje automático, percepción visual y sistemas embebidos. El Dr. Ramírez es miembro del Sistema Nacional de Investigadores de México.



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 2.5 México.

## **Development of an EEG signal analysis application through a convolution of a complex Morlet wavelet: preliminary results**

**Desarrollo de una aplicación de análisis de señales de EEG a través de una convolución de una ondícula Morlet compleja: resultados preliminares.**

José Humberto Trueba Perdomo<sup>1</sup>

[jhtrueba1@gmail.com](mailto:jhtrueba1@gmail.com)

Ignacio Herrera Aguilar<sup>1</sup>

[nacho.tecorizaba@gmail.com](mailto:nacho.tecorizaba@gmail.com)

Francesca Gasparini<sup>2</sup>

[francesca.gasparini@unimib.it](mailto:francesca.gasparini@unimib.it)

<sup>1</sup> Electronics Department, Mexico National Institute of Technology/Orizaba Institute of Technology. Oriente 9, 94320 Orizaba, Mexico

<sup>2</sup> Department of Informatics, Systems, and Communication, University of Milano Bicocca. Viale Sarca 336, 20126 Milano, Italy.



**Abstract.** This paper presents a new application for analyzing electroencephalogram (EEG) signals. The signals are pre-filtered through MATLAB's EEGLAB tool. The created application performs a convolution between the original EEG signal and a complex Morlet wavelet. As a final result, the application shows the signal power value and a spectrogram of the convoluted signal. Moreover, the created application compares different EEG channels at the same time, in a fast and straightforward way, through a time and frequency analysis. Finally, the effectiveness of the created application was demonstrated by performing an analysis of the alpha signals of healthy subjects, one signal created by the subject with eyes closed and the other, with which it was compared, was created by the same subject with eyes open. This also served to demonstrate that the power of the alpha band of the closed-eyed signal is higher than the power of the open-eyed signal.

**Keywords:** Application, Wavelet, Electroencephalogram, Signal, Analysis.

### **Desarrollo de una aplicación de análisis de señales de EEG a través de una convolución de una ondícula Morlet compleja: resultados preliminares.**

**Abstract.** Este trabajo presenta una nueva aplicación para el análisis de señales de electroencefalograma (EEG). Las señales se prefiltan a través de la herramienta EEGLAB de MATLAB. La aplicación creada genera una convolución entre la señal original del EEG y una ondícula Morlet compleja. Como resultado final, la aplicación muestra el valor de potencia de la señal y un espectrograma de la señal medida. Además, la aplicación creada compara diferentes canales de EEG al mismo tiempo, de forma rápida y sencilla, a través de un análisis de tiempo y frecuencia. Finalmente, se demostró la efectividad de la aplicación creada al realizar un análisis de las señales alfa de sujetos sanos, una señal creada por el sujeto con los ojos cerrados y la otra, con la que se comparó, fue creada por el mismo sujeto con los ojos abiertos. Esto también sirvió para demostrar que la potencia de la banda alfa de la señal de ojos cerrados es mayor que la potencia de la señal de ojos abiertos.

**Palabras clave:** Aplicación, Ondícula, Electroencefalograma, Señal, Análisis.

# 1 Introduction

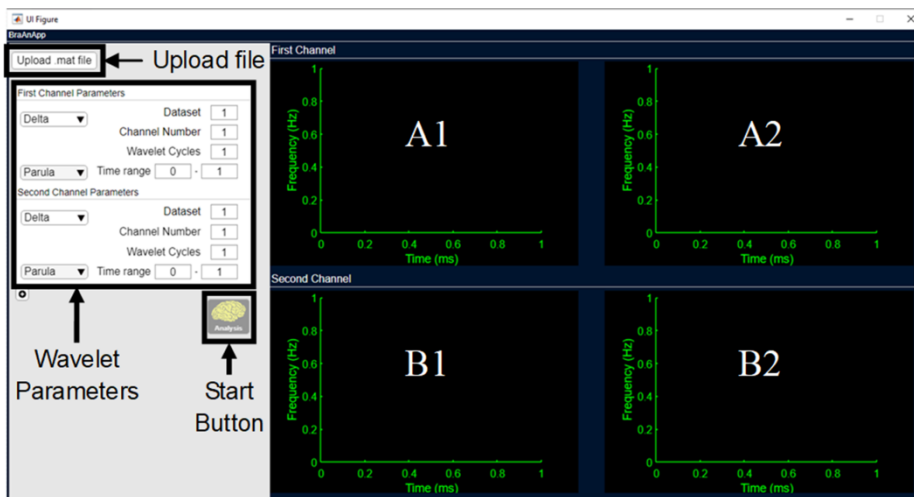
The analysis of EEG signals has become one of the most critical methods for obtaining information regarding abnormal conditions or diseases in humans, e.g., epilepsy (Adeli, Zhou, & Dadmehr, 2003), Alzheimer (Alberdi, Aztiria, & Basarab, 2016; Liu et al., 2017), and attention deficit hyperactivity disorder (Mann, Lubar, Zimmerman, Miller, & Muenchen, 1991), to name only some of them. This analysis has been achieved by obtaining the most essential neuronal information provided by brain dynamics, obtained through electrical signals produced by the brain (Omidvarnia et al., 2017). These signals have been classified into different frequency waves (delta, theta, alpha, beta, gamma) that when they present notable changes it is possible to define if the person presents some type of abnormal condition, for example, some people with Alzheimer's disease have had altered channel synchronization, loss of complexity, a slowing of frequency and synchrony or correlation between EEG signals of the different parts of the brain is reduced, which may be indicative of brain degeneration (Alberdi et al., 2016; Ieracitano, Mammone, Bramanti, Hussain, & Morabito, 2019).

In order to study and classify the characteristics of EEG signals, different mathematical algorithms have been used such as fast Fourier transform (FFT), discrete wavelet transform, autoregressive (AR) model and entropy (Zhang, Liu, Ji, & Huang, 2016), multivariate multi-scale sample entropy, support vector machines (SVM), intrinsic mode functions (IMFs) (Bhattacharyya, Pachori, & Acharya, 2017), principal component analysis (Mann et al., 1991), mentioning some of them. Of all these methods, wave transforms was selected, because without discarding any of the aforementioned, wave transforms can provide high precision of the main characteristics of EEG signals in time and frequency (Adeli et al., 2003). Moreover, it allows an analysis within a specific frequency range; because of its properties, it can decompose the original signal in its different frequency sub-bands. This method was also selected because it is possible to create a spectrogram from the waveform and obtain information by visual inspection of the signal (Adeli et al., 2003).

Within the literature analyzed to carry out this study, there has not been found an application similar to the one created, that performs an EEG signal analysis using a Morlet Wavelet Transform, in a fast and straightforward way, providing a visual (spectrogram) and numerical response. Little information was found in the literature about a quick and easy way, for analyzing and generating a spectrogram of an EEG signal. Therefore, in this paper, an application that develops an analysis by performing a convolution is presented (Fig. 1). The left command window has a button at the top

to upload a .mat file, also has a panel to insert the parameters of the wavelet within which are the number of cycles and an option to select the frequency band to analyze. In addition, the user is allowed to indicate the dataset number and the EEG channel number that will be analyzed. Finally, on the same left command window the button to start the analysis is found. The graphical spaces A1 and A2 show the spectrograms of the first channel, meanwhile spaces B1 and B2 show the spectrograms of the second channel. The comparison between both channels on the rank frequencies from 0 to 40 Hz can be seen in the right side figures. The panels on the left (1) show the zoom of a frequency band from each panel on the right (2), respectively.

The original signal and the complex Morlet wavelet are convoluted to generate the spectrogram of an EEG signal. Also, the created application calculates the power of each frequency band of the EEG signal to provide more accurate results.

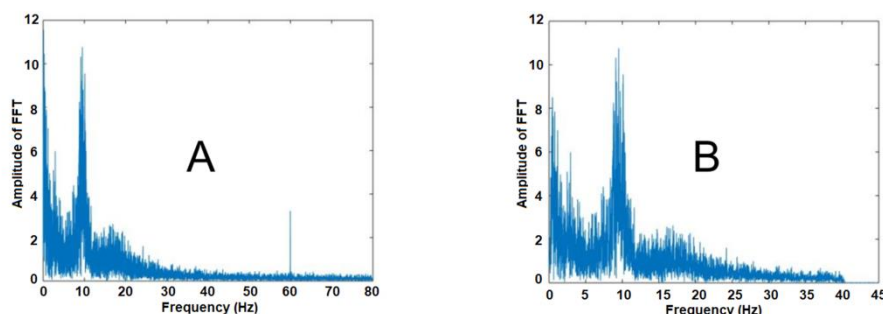


**Figure 1.** Main application interface created with MATLAB. The left command window has a button at the top to upload a file, a panel to insert the parameters of the wavelet and a button to start the analysis. The graphical spaces A1 and A2 show the spectrogram of the first channel. Spaces B1 and B2 show the spectrogram of the second channel.

## 2 Methods

### Signal pre-processing

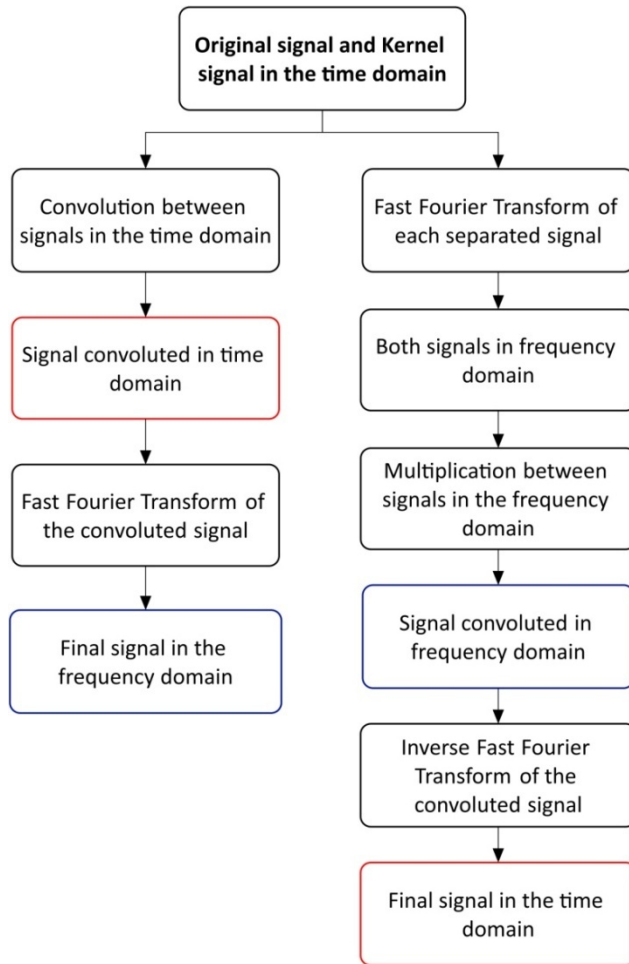
Because the program does not filter the data, the data must be previously filtered with the EEGLAB toolbox provided by MATLAB or another filter (Fig. 2). For this study, the power of the alpha band (8 - 14 Hz (Lobier, Palva, & Palva, 2017)) will be analyzed, therefore a high-pass filter at 0.5 Hz is useful and recommended to minimize slow drifts (X Cohen, 2014), and 40 as the edge of the low-pass filter can help to remove high-frequency artifacts (X Cohen, 2014), it will also help to attenuate electrical line noise that occurs at 60 Hz in Mexico and 50 Hz in Europe produced by the power supply current. EEGLAB works with the created application, saving the most relevant information of the EEG signals that will be used to run the application. For example, the location of the electrodes must be saved with the number of each one, not just with the letter (e.g., Fz) in order that the application will know the number of the channel to analyze.



**Figure 2.** A) the Fast Fourier Transform of the unfiltered Pz channel signal, B) EEG signal but after applying a 0.5 to 40 Hz filter.

### Algorithm

This program performs a convolution between EEG signals and a Complex Morlet Wavelet (Fig. 3), an algorithm described in (X Cohen, 2014).



**Figure 3.** Flowchart of the interchangeability of convolution in the time domain and frequency domain multiplication. The signals resulting from the processes shown within the blue box are the same, as are the signals resulting from the processes within the red box (X Cohen, 2014); using a shorter computation time the cycle on the right side of the flow chart.

The convolution method was chosen together with a wavelet transform, since the FFT method does not allow to visualize an analysis based on the time domain, whereas with the wavelet cycles a decomposition of the data in the frequency domain, in the time domain or both, can be performed (Pattnaik, Dash, & Sabut, 2016; X Cohen, 2014). Due to the limitations that the wavelets present as a band-pass filter, it was necessary to use a complex Morlet wavelet described in (X Cohen, 2014), which is created by multiplying, point by point, a Gaussian window and a complex sine wave:

$$\text{Complex Morlet Wavelet (CMW)} = Ae^{-t^2/2s^2} \cdot e^{i2\pi ft}$$

(1)

with  $A$  being a frequency band-specific scaling factor,  $e$  a constant called Euler's number,  $t$  is time,  $i$  is the imaginary operator (X Cohen, 2014),  $f$  is frequency (Hz) and  $s$  is the standard deviation of the Gaussian:

$$s = \frac{n}{2\pi f}$$

(2)

where  $n$  is the number of wavelet cycles.

Also, this wavelet will serve to extract power information from the EEG data (X Cohen, 2014). Furthermore, the result of convolution in the time domain can be represented by a multiplication between two vectors in the frequency domain. To achieve this, the FFT of the original signal and the CMW (kernel) is calculated separately. Finally, both resulting FFT signals are multiplied and the IFFT of the convoluted signal is calculated:

$$\text{Final signal} = \text{IFFT} (\text{FFT} (\text{original signal}) \cdot \text{FFT} (\text{kernel}))$$

(3)

For this study, multiplication in the frequency domain was chosen because it reduces the computing time (X Cohen, 2014) of MATLAB. It is important to mention that the created application uses automatically the frames per epoch and the trials of each dataset to execute the convolution, so the result will be the same as the length of the kernel plus the length of the original signal minus one, as stated in (X Cohen, 2014).

After convolution, the real part of the final signal is introduced into a matrix, and a filled contour plot is created with the isolines of the final signal power values. In which it will be observed in the axis of the ordinate the range of the frequency bands, from 1 up to 40 Hz; in the axis of the abscissa, it will be appreciated the time, that in this case are milliseconds.

## **Database**

An EEG motor movement database (Schalk, Mcfarland, Hinterberger, Birbaumer, & Wolpaw, 2004) was used for evaluating the performance of the application. The used database consists of one hundred nine subjects performed different experiments, including opening and closing the fists of the hands, both physically and imaginatively, in order to distinguish these movements through the EEG signals (Goldberger et al., 2000). These signals were taken from an EEG of 64 channels, placed on the scalp according to the 10-20 international system and, each signal was sampled at 160 samples per second. The BCI2000 system was used to acquire the signals, developed and tested in (Schalk et al., 2004), which incorporates any brain signals, signal processing methods, output devices, and operating protocols. However, for this study, only the two baselines of the study were taken into account, which were created by the subject keeping eyes closed (without motor movement) for one minute and the other baseline keeping eyes open, for one minute as well.

Moreover, at a later stage, different analyses will be performed taking into account the potentials evoked due to imaginary movements, as well as the analysis of other mental states, in order to validate the application and demonstrate that it can be used as part of a multimodal system in order to find unique biomarkers that serve for the early diagnosis of different diseases such as Alzheimer's, autism, tumours, etc (Alberdi et al., 2016; Siuly & Zhang, 2016).

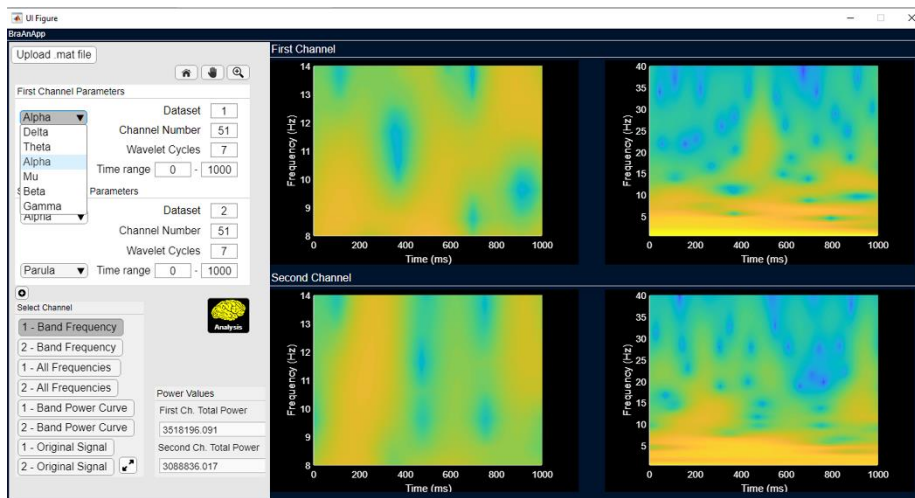
### 3 Results

The created application compares different EEG channels at the same time, in a fast and straightforward way, through a time and frequency analysis. The baselines of the EEG database were selected with the idea of demonstrating that the alpha waves have a higher power when the subject has closed eyes than when the subject has open eyes, this following the established in different previous studies (Butler & Glass, 1974; Kan, Croarkin, Phang, & Lee, 2017; Morgan, McDonald, & Macdonald, 1971; Teplan, 2002; Valipour, Shaligram, & Kulkarni, 2013). Moreover, the PZ (51), CZ (11), FZ (34), and O2 (63) electrodes were selected, because they are commonly used since they have given significant results (Hussain et al., 2017; Rushby et al., 2016; Wyckoff, Sherlin, Ford, & Dalke, 2015) in terms of the power of the alpha band. Just to demonstrate the functionality of the application, 3 random subjects were taken into consideration to verify that the power of the alpha band is greater when moving the electrode from the front to the parietal lobe (Valipour et al., 2013).

The created application asks the user to select the necessary characteristics for the analysis of EEG signals, i.e.: dataset number, channel number, number of wavelet cycles, the time range to be displayed on the plot (the application generates the analysis using the full-time range, not just the one selected by the user). Also, the colormap can be selected: Parula or Jet. Both colormaps have been proposed because colormaps in which their brightness function is linear, such as Parula (the default since MATLAB version R2014b), are still subject to usability problems (Helfman, 2015). However, the aim is for future researchers to be able to test both colormaps and have more reliable results. By the other hand, it is important to declare that the total number of wavelet frequencies used in the created application is 40, in a range from 0.5 to 40 Hz, since the frequencies of interest are within this range and the wave frequencies cannot be above the Nyquist frequency (half the sampling frequency) (X Cohen, 2014). The created application uses the sampling rate (Hz) of each dataset to select the center of the wavelets as a zero.

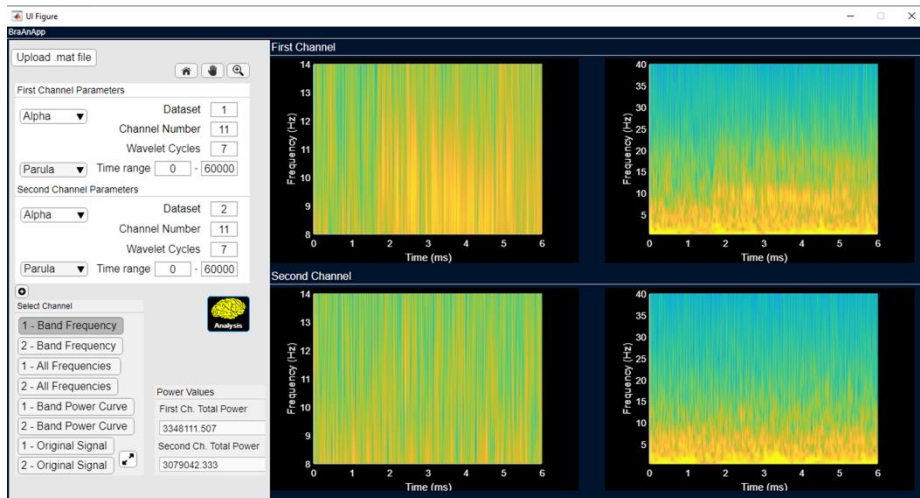
Figure 4 shows the Pz (51) channel spectrograms after having performed an analysis using seven cycles for the wavelet and selecting a time from 0 to 1000 ms to display on the plot. Dataset number 1 was defined by the baseline with the eyes closed, while the number 2 was created while the eyes were kept open, both datasets belong to the test subject number 1. The color limits for these graphics are set by the range values of the frequency band selected by the user (see Figure 4). Therefore, the areas with the highest power within the same frequency band can be identified.





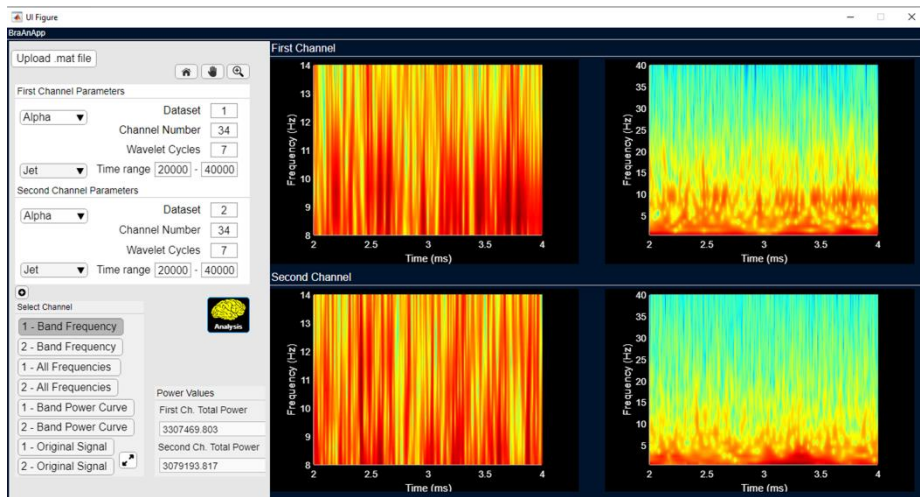
**Figure 4.** Analysis of channel Pz (51). Time range selected to display from 0 to 1000 ms. Both spectrograms at the top belong to the dataset of subject 1 with closed eyes. The two spectrograms at the bottom belong to the dataset of subject 1 with open eyes. On the left, it can be seen the drop-down menu with the six frequency bands which are selectable for different analyses.

Figure 5 shows the results of the Cz (11) channel analysis. For this analysis, the full-time range of the signal (60000 ms) was selected to be displayed on the plot. It can be seen that the dataset with closed eyes has the power of the alpha band higher than that of open eyes. However, in the lower part of the section where the user specifies the characteristics for analysis, the application provides a result with the value of the power of the band that the user has previously selected, in this case, the values shown belong to the alpha band.



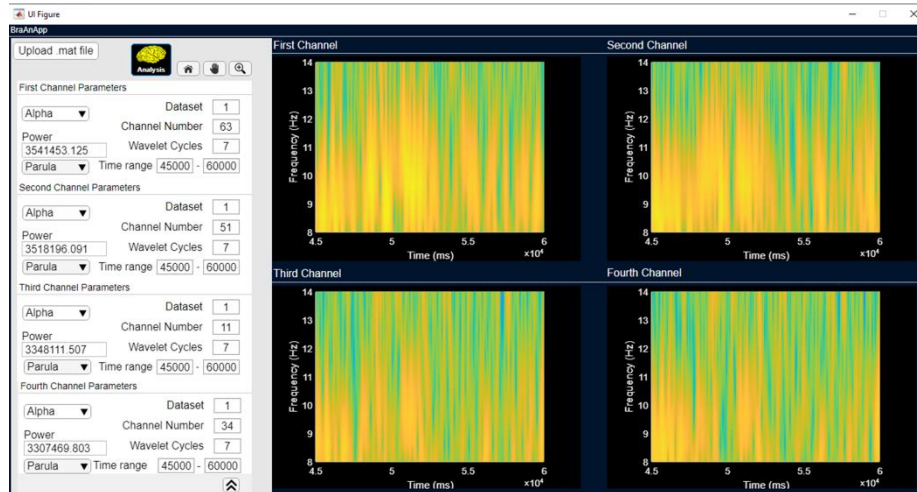
**Figure 5.** Analysis of channel Cz (11). Time range selected to display from 0 to 60000 ms. Both spectrograms at the top belong to the dataset of subject 1 with closed eyes. The two spectrograms at the bottom belong to the dataset of subject 1 with open eyes.

Figure 6 shows the spectrograms of channel Fz (34). Again it is observed that the dataset with closed eyes (1) has higher power than the dataset with open eyes (2). For this analysis, seven cycles were also used for the wavelet and the results from 20000 to 40000 ms were shown. Also, the colormap jet was selected only as a demonstration object for the reader's knowledge.



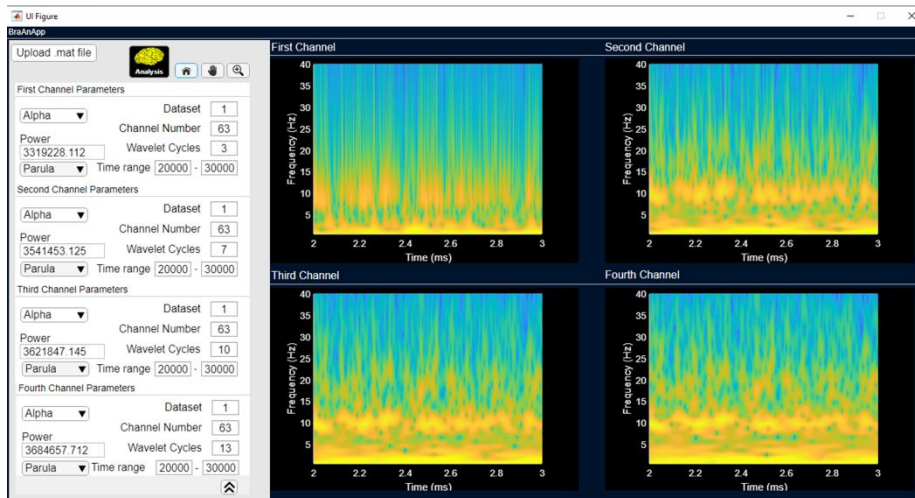
**Figure 6.** Analysis of channel Fz (34). Time range selected to display from 20000 to 40000 ms. Both spectrograms at the top belong to the dataset of subject 1 with closed eyes. The two spectrograms at the bottom belong to the dataset of subject 1 with open eyes.

Figure 7 shows another of the capabilities of the created application, which is to compare four channels at the same time. In this case, the channels O2 (63), PZ (51), CZ (11), and FZ (34) were compared. It can be seen from each of the spectrograms that the power is more significant as we move closer to the electrodes at the back of the brain, following the conclusions stipulated in [14]. Also, to have a more precise response, the application created shows the total value of the power of the EEG band previously selected for each of the analyzed channels ( $\mu\text{V}$ ).



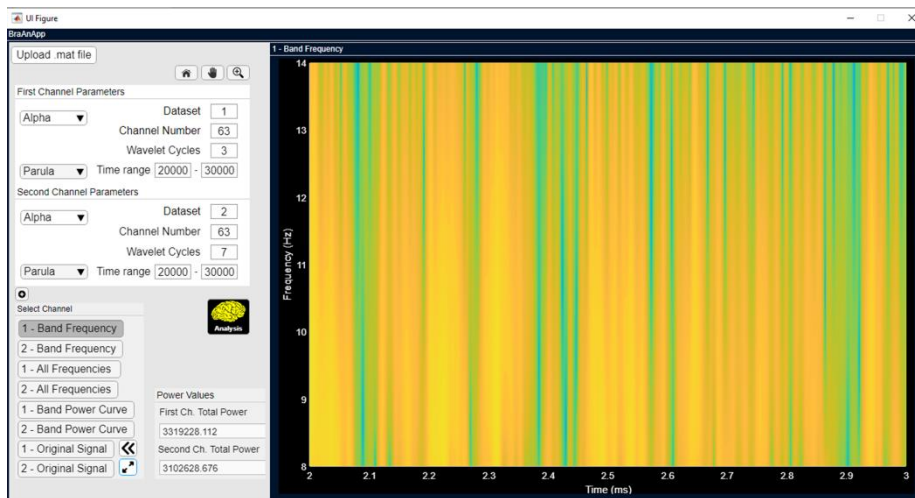
**Figure 7.** Analysis of O2 (top left), Pz (top right), Cz (bottom left), and Fz (bottom right) channels. Time range selected to display from 45000 to 60000 ms. Frequency range from 8 to 14 (alpha wave).

The number of cycles of the wavelet can be changed and defined every time before the analysis performed by the application. Figure 8 presents an analysis performed on the O2 (63) channel, increasing the cycles for the wavelet to demonstrate the differences produced by these parameters in each analysis. In order to achieve better temporal precision using fewer cycles (e.g., 3), or if greater frequency precision is required more cycles can be selected (e.g., 10). Finally, an average analysis can be performed by selecting an intermediate number of cycles (e.g., 7).

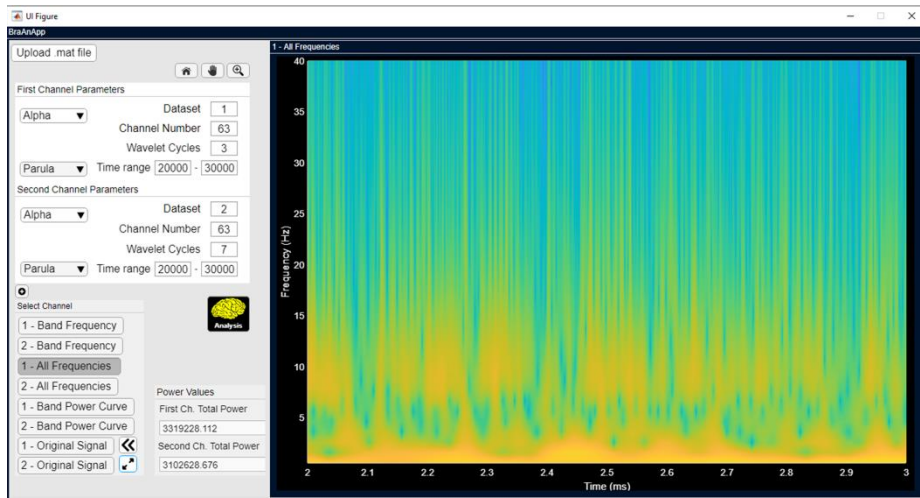


**Figure 8.** Analysis of channel O2 (63). Time range selected to display from 20000 to 30000 ms. Different numbers of wavelets cycles selected: 3 (top left), 7 (top right), 10 (bottom left), and 13 (bottom right).

Figure 9 shows how the created application can enlarge the spectrograms for more accurate analysis. Also, it is possible to observe each of the curves of the power bands created after convolution (Fig. 10), taking into account only the values provided by the band selected in the drop-down menu. Finally, if the user wishes to observe the original signal, the application allows this by selecting one of the last options from the bottom panel on the left side of the application (Fig. 11).

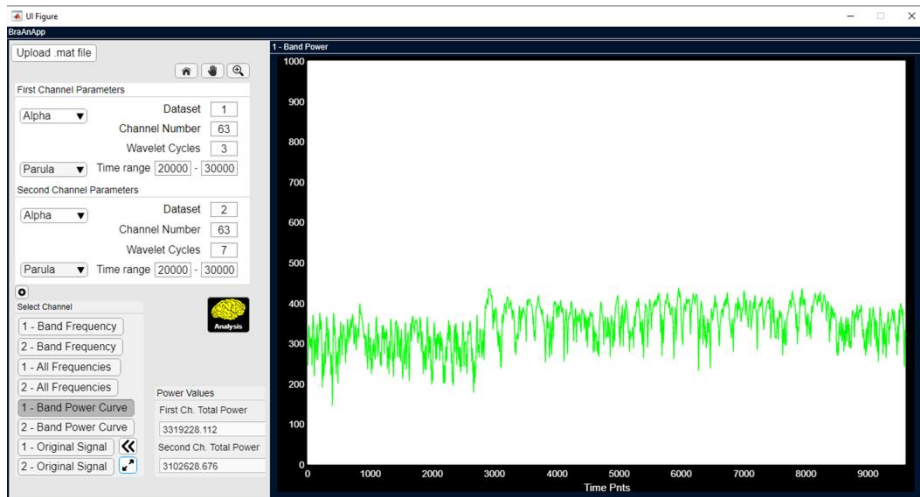


A)



B)

**Figure 9.** A) Maximization of the spectrogram of the frequency band selected (alpha) of the first channel. B) Maximization of the spectrogram of all frequencies of the first channel.



**Figure 10.** Signal created after convolution, taking into account only the values present within the range of the frequency band selected by the user.

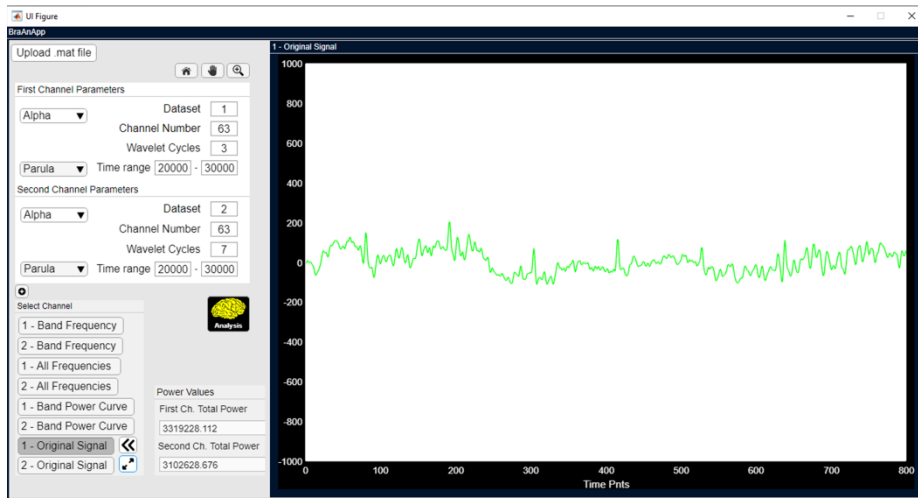


Figure 11. Original EEG signal from channel O2 (63).

## 4 Conclusions

A new application was described in this paper, which performs a convolution between the signals of an EEG and a complex Morlet wavelet, an algorithm described in (X Cohen, 2014). In addition, the application shows the spectrogram of each EEG signal and is able to calculate the power of each frequency band of the EEG signal. The created application allows the user to select the basic characteristics needed for an EEG signal analysis and to select the EEG channel number, in order to compare different EEG channels at the same time.

### Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## 5 References

- Adeli, H., Zhou, Z., & Dadmehr, N. (2003). Analysis of EEG records in an epileptic patient using wavelet transform. *Journal of Neuroscience Methods*, 123(1), 69–87. [https://doi.org/https://doi.org/10.1016/S0165-0270\(02\)00340-0](https://doi.org/https://doi.org/10.1016/S0165-0270(02)00340-0)
- Alberdi, A., Aztiria, A., & Basarab, A. (2016). On the early diagnosis of Alzheimer ' s Disease from multimodal signals : A survey. *Artificial Intelligence In Medicine*, 71, 1–29. <https://doi.org/10.1016/j.artmed.2016.06.003>
- Bhattacharyya, A., Pachori, R. B., & Acharya, U. R. (2017). Tunable-Q wavelet transform based multivariate sub-band fuzzy entropy with application to focal EEG signal analysis. *Entropy*, 19(3), 99. <https://doi.org/10.3390/e19030099>
- Butler, S. R., & Glass, A. (1974). Asymmetries in the electroencephalogram associated with cerebral dominance. *Electroencephalography and Clinical Neurophysiology*, 36, 481–491. [https://doi.org/https://doi.org/10.1016/0013-4694\(74\)90205-3](https://doi.org/https://doi.org/10.1016/0013-4694(74)90205-3)
- Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P., Mark, R., ... Stanley, H. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101(23), 215–220. <https://doi.org/https://doi.org/10.1161/01.CIR.101.23.e215>
- Helfman, J. (2015). Hue tinting for interactive data visualization. *Human Vision and Electronic Imaging XX. Society of Photo-Optical Instrumentation Engineers (SPIE)*, 9394, 1–12. <https://doi.org/10.1117/12.2079543>
- Hussain, L., Aziz, W., Alowibdi, J. S., Habib, N., Rafique, M., Saeed, S., & Kazmi, S. Z. H. (2017). *Symbolic time series analysis of electroencephalographic (EEG) epileptic seizure and brain dynamics with eye-open and eye-closed subjects during resting states*. 36(21), 1–12. <https://doi.org/10.1186/s40101-017-0136-8>
- Ieracitano, C., Mammone, N., Bramanti, A., Hussain, A., & Morabito, F. C. (2019). A convolutional neural network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings. *Neurocomputing*, 323, 96–107. <https://doi.org/10.1016/j.neucom.2018.09.071>
- Kan, D. P. X., Croarkin, P. E., Phang, C. K., & Lee, P. F. (2017). EEG differences between eyes-closed and eyes-open conditions at the resting stage for euthymic participants. *Neurophysiology*, 49(6), 432–440. <https://doi.org/10.1007/s11062-018-9706-6>

Liu, J., Li, M., Pan, Y., Lan, W., Zheng, R., Wu, F., & Wang, J. (2017). Complex brain network analysis and its applications to brain disorders : A survey. *Complexity*, 2017, 1–27.

Lobier, M., Palva, J. M., & Palva, S. (2017). High-alpha band synchronization across frontal, parietal and visual cortex mediates behavioral and neuronal effects of visuospatial attention. *NeuroImage*, 165, 222–237. <https://doi.org/10.1016/j.neuroimage.2017.10.044>

Mann, C. A., Lubar, J. F., Zimmerman, A. W., Miller, C. A., & Muenchen, R. A. (1991). Quantitative analysis of EEG in boys with attention-deficit-hyperactivity disorder : Controlled study with clinical implications. *Pediatric Neurology*, 8(1), 30–36. [https://doi.org/https://doi.org/10.1016/0887-8994\(92\)90049-5](https://doi.org/https://doi.org/10.1016/0887-8994(92)90049-5)

Morgan, A. H., Mcdonald, P. J., & Macdonald, H. (1971). Differences in bilateral alpha activity as a function of experimental task, with a note on lateral eye movements and hypnotizability. *Neuropsychologia*, 9(4), 459–469. [https://doi.org/https://doi.org/10.1016/0028-3932\(71\)90011-X](https://doi.org/https://doi.org/10.1016/0028-3932(71)90011-X)

Omidvarnia, A., Pedersen, M., Vaughan, D. N., Walz, J. M., Abbott, D. F., Zalesky, A., & Jackson, G. D. (2017). Dynamic coupling between fMRI local connectivity and interictal EEG in focal epilepsy : A wavelet analysis approach. *Human Brain Mapping*, 38(11), 5356–5374. <https://doi.org/10.1002/hbm.23723>

Pattnaik, S., Dash, M., & Sabut, S. K. (2016). DWT-based feature extraction and classification for motor imaginary EEG signals. *Proceedings of 2016 International Conference on Systems in Medicine and Biology*, (January), 186–191. <https://doi.org/10.1109/ICSMB.2016.7915118>

Rushby, J. A., Mcdonald, S., Fisher, A. C., Kornfeld, E. J., Blasio, F. M. De, Parks, N., & Piguet, O. (2016). Brain volume loss contributes to arousal and empathy dysregulation following severe traumatic brain injury. *NeuroImage: Clinical*, 12, 607–614. <https://doi.org/10.1016/j.nicl.2016.09.017>

Schalk, G., Mcfarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000 : A General-Purpose Brain-Computer Interface ( BCI ) System. *IEEE Transactions on Biomedical Engineering*, 51(6), 1034–1043. <https://doi.org/10.1109/TBME.2004.827072>

Siuly, S., & Zhang, Y. (2016). Medical big data : neurological diseases diagnosis through medical data analysis. *Data Science and Engineering*, 1(2), 54–64. <https://doi.org/10.1007/s41019-016-0011-3>



Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review*, 2(2), 1–11.

Valipour, S., Shaligram, A. D., & Kulkarni, G. R. (2013). Spectral analysis of EEG signal for detection of alpha rhythm with open and closed eyes. *International Journal of Engineering and Innovative Technology (IJEIT)*, 3(6), 1–4.

Wyckoff, S. N., Sherlin, L. H., Ford, N. L., & Dalke, D. (2015). Validation of a wireless dry electrode system for electroencephalography. *Journal of NeuroEngineering and Rehabilitation*, 12(1), 1–9. <https://doi.org/10.1186/s12984-015-0089-2>

X Cohen, M. (2014). *Analyzing neural time series data: Theory and practice*. London, UK: The MIT Press.

Zhang, Y., Liu, B., Ji, X., & Huang, D. (2016). Classification of EEG signals based on autoregressive model and wavelet packet decomposition. *Neural Process Lett*, 45(2), 365–378. <https://doi.org/10.1007/s11063-016-9530-1>

## Author biography

**José Humberto Trueba Perdomo** is a Master student in the Department of Electronics, Technological Institute of Orizaba, Mexico. His research interests include pain analysis through biometric signals.

**Ignacio Herrera Aguilar** is the Director of the Signal Processing Laboratory at the Technological Institute of Orizaba, Mexico. He has a PhD in Automatic Systems and a MSc in Electronic Engineering. He specializes in Control and Automation of robots and mechanisms that can be used in the industry, as well as in motor rehabilitation systems for human beings.

**Francesca Gasparini** is an associate professor at the University of Milano Bicocca (Italy). She is head of the Multimedia Signal Processing Laboratory. Her research activity is concentrated in multimedia signal processing. She is particularly interested in processing and analysis of neurophysiological data as well as psychophysical ones, using machine learning techniques, genetic programming and optimization algorithms.



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 2.5 México.