

Centro Nacional de Investigación y Desarrollo Tecnológico

Subdirección Académica

Departamento de Ciencias Computacionales

TESIS DE MAESTRÍA EN CIENCIAS

**Desarrollo de un sistema interactivo basado en visión artificial
para la rehabilitación del tobillo**

presentada por
Ing. Ana Karen Gama Velasco

como requisito para la obtención del grado de
Maestra en Ciencias de la Computación

Directora de tesis
Dra. Andrea Magadán Salazar

Codirector de tesis
Dr. Andrés Blanco Ortega

Cuernavaca, Morelos a 21 de febrero del 2018
OFICIO No. DCC/040/2018

Asunto: Aceptación de documento de tesis

C. DR. GERARDO V. GUERRERO RAMÍREZ
SUBDIRECTOR ACADÉMICO
PRESENTE

Por este conducto, los integrantes de Comité Tutorial de la **Ing. Ana Karen Gama Velasco**, con número de control M15CE079, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis profesional titulado **"Desarrollo de un sistema interactivo basado en visión artificial para la rehabilitación de tobillo"** y hemos encontrado que se han realizado todas las correcciones y observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

DIRECTORA DE TESIS

Dra. Andrea Macedán Salazar
Doctorado en Ciencias Computacionales
10664097

CO-DIRECTOR DE TESIS

Dr. Andrés Blanco Ortega
Doctor en Ciencias en Ingeniería Eléctrica
6559298

REVISOR 1

Dr. Raúl Pinto Elías
Doctor en Ciencias en la Especialidad de
Ingeniería Eléctrica
3890453

REVISOR 2

M.C. Gerardo Reyes Salgado
Maestro en Ciencias de la Computación
2493370

C.c. M.C. María Elena Gómez Torres - Jefa del Departamento de Servicios Escolares.
Estudiante
Expediente

NACS/ma



Cuernavaca, Mor., 23 de febrero de 2018
OFICIO No. SAC/128/2018

Asunto: Autorización de impresión de tesis

ING. ANA KAREN GAMA VELASCO
CANDIDATA AL GRADO DE MAESTRA EN CIENCIAS
DE LA COMPUTACIÓN
PRESENTE

Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado "**Desarrollo de un sistema interactivo basado en visión artificial para la rehabilitación de tobillo**", ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

ATENTAMENTE
"CONOCIMIENTO Y TECNOLOGÍA AL SERVICIO DE MÉXICO"

DR. GERARDO VICENTE GUERRERO RAMÍREZ
SUBDIRECTOR ACADÉMICO



SEP TecNM
CENTRO NACIONAL
DE INVESTIGACIÓN
Y DESARROLLO
TECNOLÓGICO
SUBDIRECCIÓN
ACADÉMICA

Cp. M.T. María Elena Gómez Torres - Jefa del Departamento de Servicios Escolares.
Expediente

GVGR/mcr



Le dedico este trabajo a:

Mi familia y a mi novio Leopoldo

Gracias por su apoyo incondicional.

Agradecimientos:

Primeramente agradezco a Dios que me ha brindado la salud y permitido culminar mis estudios de maestría. Gracias por las pruebas que me hacen crecer como persona, me permiten dar lo mejor de mí y me acercan más a ti.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) y al Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), por el apoyo brindado para la realización del presente trabajo.

A mis padres Ana Velasco y Juan Gama por haberme dado la vida, y enseñarme que las metas son alcanzables. Gracias por orientarme y brindarme la oportunidad de estudiar siguiendo su ejemplo día con día, honrándolos con mis éxitos que ahora también son suyos. También agradezco que me hayan dado lo más maravilloso, mis hermanos, de los cuales me siento muy orgullosa. Los quiero mucho.

Agradezco a la Dra. Andrea Magadán Salazar quien me ha guiado durante todo este proceso, muchas gracias por todo el tiempo, asesorías, enseñanzas y consejos que me ha brindado a lo largo de este trayecto, es una gran persona y una excelente asesora.

Al Dr. Andrés Blanco Ortega quien codirigió esta tesis, agradezco de igual manera todas las enseñanzas y recomendaciones que me ha proporcionado, también por haberme brindado la oportunidad de participar en el concurso ENEIT, el cual me ha dejado múltiples experiencias que con orgullo puedo compartir.

Agradezco a mi comité revisor: al Dr. Raúl Pinto Elías, al Dr. Gerardo Reyes Salgado y al Dr. José Ruíz Asencio, por sus comentarios, observaciones, recomendaciones y el tiempo dedicado a la revisión del presente trabajo.

A los especialistas Lic. Guillermo Gutiérrez Vázquez (Licenciado en terapia física) y al Lic. Eduardo Sánchez Rivero (Licenciado en Psicología) por su colaboración en el presente trabajo.

Al Ing. Oscar Fernando Oliveros Riego y al Ing. Agustín Erasmo Juárez Martínez, les agradezco por su colaboración y apoyo con la implementación del rehabilitador.

Muchas gracias a la Dra. Olivia, al Dr. Alcantara, al Dr. Marco Oliver †, al Dr. Lavallo, al Dr. Salomón y cada uno de mis profesores de Cenidet, por su dedicación, atención y su tiempo.

Itzel Blanco sabes que tú eres mi hermana de corazón, agradezco tener personas como tú en mi vida, toda la ayuda y el apoyo que me das y por ser una gran persona

A Arturo Pérez y Charles Velázquez, quienes me apoyaron siempre que lo necesité, muchas gracias por sus consejos y recomendaciones.

A mis compañeros de posgrado Diana, Pedro, Nelva, Marianita, Anahí, Luis, Socrates, Lupita, Javier, Andrea, Michel, Manuel, Jessy, Edgar, Carlos y a todos aquellos que me preguntaban todo el tiempo que si ya iba a terminar, por su amistad, por todos los momentos de alegría y por hacer de mi estancia en Cenidet una experiencia formidable.

Por último, a Leopoldo Alfonso Guillén, gracias por tu apoyo incondicional, por tu paciencia, por estar siempre dispuesto a ayudarme y la motivación que me das día con día. Mi vida es más feliz a tu lado y sé que puedo contar contigo en cualquier situación por más difícil que sea al igual que tú puedes confiar en mí. Te amo.

Ana Karen Gama Velasco

Resumen

Los seres humanos están expuestos a sufrir accidentes traumáticos que pueden causar daños de manera permanente. Por lo general, éstos son causados por exceso de esfuerzo o por un golpe, impidiendo que se puedan llevar a cabo actividades de la vida cotidiana.

Dichos incidentes traumáticos requieren de una rehabilitación para recuperar la fuerza y movilidad perdida; pero no siempre la recuperación que otorga una clínica es la más adecuada ya que cuentan con poco personal y equipo.

Por otro lado, el proceso de rehabilitación requiere de varios especialistas y varias sesiones de rehabilitación, haciendo que sea un proceso muy costoso.

En esta tesis se desarrolló un sistema inteligente de rehabilitación de tobillo conformado por un rehabilitador de tobillo de 1 grado de libertad, un juego serio y un sistema de visión artificial que realiza el reconocimiento de expresiones faciales para conocer la motivación o molestia del paciente con el equipo de rehabilitación.

El sistema de visión artificial fue desarrollado en visual *c++*, utilizando para el seguimiento del rostro y sus componentes la librería *facetracking*, para el reconocimiento de expresiones faciales se evaluaron tres algoritmos, de los cuales se obtuvo el mejor desempeño con J48. El sistema desarrollado propone la detección de tres expresiones faciales diferentes a las básicas, que ayudan a monitorear al paciente, de tal forma que cuente con información sobre la rehabilitación, como por ejemplo, si tuvo alguna molestia, el nivel de movimiento realizado y el tiempo dedicado a su rehabilitación. El juego serio fue implementado en el motor de videojuegos *unity engine*. Dichos sistemas fueron integrados a un rehabilitador de tobillo de un grado de libertad controlado mediante *arduino*, con el objetivo de retroalimentarlo. Por otro lado, la dificultad del juego, que el paciente debe realizar, considera el movimiento y fuerza que la recuperación que él mismo requiere. Este sistema permite que este tipo de rehabilitación pueda llevarse a cabo de manera amena, tanto en la clínica como en el hogar.

Índice

Capítulo 1: Introducción	1
1.1 Problemática	1
1.2 Objetivos	3
1.2.1 Objetivo general	3
1.2.2 Objetivos específicos:	3
1.3 Alcances y limitaciones	4
1.3.1 Alcances	4
1.3.2 Limitaciones	4
1.4 Justificación	4
1.5 Organización de la tesis	5
Capítulo 2: Reconocimiento automático de expresiones faciales.....	6
2.1 Estado del arte referente a FER.....	7
2.2 Análisis del estado del arte referente a FER	17
2.3 Sensor Kinect.....	19
2.4 Expresiones faciales propuestas.....	22
2.5 Esquema de funcionamiento del sistema de expresiones faciales	23
2.6 Creación de las bases de datos	24
2.7 Clasificación	26
2.7.1 Naïve Bayes [24].....	26
2.7.2 Máquinas de soporte vectorial.....	27
2.7.3 J48.....	29
2.8 Herramientas para clasificación	30
2.8.1 Weka.....	30
2.8.2 RStudio	31
2.9 Discusión	31
Capítulo 3: Juegos serios	32
3.1 Estado del arte referente a juegos serios	32
3.2 Análisis del estado del arte referente a juegos serios.....	41
3.3 Diseño y desarrollo del juego serio	43
3.3.1 Diseño del rehabilitador del tobillo de 1 gdl	44
3.3.2 Psicología del color en juegos serios.....	46

3.3.3 Animaciones	47
3.4 Discusión	48
Capítulo 4: Diseño de la integración	49
4.1 Discusión.....	52
Capítulo 5: Experimentación	53
5.1 Ambiente de pruebas	53
5.2 Plan de pruebas	53
5.2.1 Prueba 1: Condiciones para el uso de Kinect	54
5.2.2 Prueba 2: Rango de movimiento de tobillo en personas sanas	56
5.2.3 Prueba 3: Comparación de algoritmos	57
5.2.4 Prueba 4: Funcionalidad de los sistemas integrados	59
5.2.5 Prueba 5: Aplicar modelo de clasificación con los nuevos datos	62
5.3 Discusión.....	63
Capítulo 6: Conclusiones	64
6.1 Objetivos alcanzados.....	64
6.2 Aportaciones.....	65
6.3 Trabajos futuros	66
6.4 Productos adicionales.....	66
6.5 Conclusiones.....	66
Referencias.....	68
Anexos.....	73
ANEXO I: Bases de datos para reconocimiento de expresiones faciales	74

Índice de figuras

Figura 1.1 Movimiento de dorsiflexión-plantarflexión.	2
Figura 2.1 Factores que influyen en las expresiones faciales [6].	7
Figura 2.2 Metodología MSV[7].	8
Figura 2.3 Etapas FER.	16
Figura 2.4 Sensor Kinect.	19
Figura 2.5 Unidades de acción identificadas con Kinect.	20
Figura 2.6 Movimientos de cabeza identificados con Kinect.	20
Figura 2.7 Candide-3.	22
Figura 2.8 Vector de características.	22
Figura 2.9 Esquema de funcionamiento para la etapa de entrenamiento del sistema de reconocimiento de expresiones faciales.	24
Figura 2.10 Esquema de funcionamiento para validar los modelos de clasificación para el sistema de reconocimiento de expresiones faciales (Etapa de prueba).	24
Figura 2.11 Almacenamiento.	25
Figura 2.12 Elementos MSV [25].	28
Figura 2.13 Función kernel RBF.	29
Figura 3.1 Estructura de los niveles del juego serio.	43
Figura 3.2 Diseño del robot paralelo [45].	44
Figura 3.3 Tarjeta arduino Mega.	45
Figura 3.4 Sensor de fuerza.	45
Figura 3.5 Prototipo físico.	46
Figura 3.6 Psicología del color.	46
Figura 3.7 Diagrama de estados.	47
Figura 3.8 Personajes.	48
Figura 4.1 Esquema de la integración.	51
Figura 5.1 Comparación de condiciones de luz.	54
Figura 5.2 Condiciones de distancia.	55
Figura 5.3 Respuesta de la máscara candide-3 ante la presencia de accesorios.	56
Figura 5.4 Rango de movimiento en personas sanas.	56
Figura 5.5 Modelo J48 generado.	59
Figura 5.6 Recepción de señal de fuerza.	59
Figura 5.7 Pruebas con el sistema integrado.	60
Figura 5.8 Paciente motivado en el nivel 1.	61
Figura 5.9 Paciente desmotivado nivel 1.	61
Figura 5.10 Paciente adolorido en nivel 1.	62

Índice de Tablas

Tabla 2.1 Expresiones faciales básicas.	16
Tabla 2.2 Experimentación survey de FER.	17
Tabla 2.3 Resumen del estado del arte FER.	18
Tabla 2.4 Comparación de unidades de acción.	21
Tabla 2.5 Comparación de unidades de acción en las expresiones faciales básicas.	21
Tabla 2.6 Descripción de las clases propuestas.	23
Tabla 2.7 Ejemplo de selección.	26
Tabla 3.1 Resumen del estado del arte referente a juegos serios.	42
Tabla 3.2 Variables.	47
Tabla 4.1 Herramientas de desarrollo.	50
Tabla 4.2 Valores considerados para la integración.	52
Tabla 5.1 Rango de movimiento para cada nivel del juego serio.	57
Tabla 5.2 Comparación de algoritmos en fase de entrenamiento.	57
Tabla 5.3 Comparación de algoritmos en fase de pruebas.	58
Tabla 6.1 Resultados de los objetivos específicos.	65

Glosario

- **AAM:** Por sus siglas en inglés *Active Apperence Model*, es un algoritmo estadístico de visión artificial que busca recuperar una descripción paramétrica de un determinado objeto.
- **Candide-3:** Malla parametrizada desarrollada específicamente para la codificación basada en el modelo de rostros humanos, la cual contiene 113 vértices y 168 superficies.
- **CNN:** Redes neuronales convolucionales, por sus siglas inglés *Convolutional Neural Network*, son una clase de redes neuronales avanzadas aplicadas al análisis de imágenes visuales.
- **CSV:** del inglés *Comma-separeted values*, son un tipo de documento en formato abierto sencillo para representar datos en forma de tabla.
- **Dorsiflexión:** Acción de la articulación del tobillo en la que el pie se dirige hacia arriba, en un rango máximo de 25°.
- **Facetracking:** Librería del *sdk* de Kinect para la detección y seguimiento del rostro.
- **FACS:** Por sus siglas en inglés *Facial Action Code System*, es el sistema de codificación facial donde los movimientos de la cara tienen asignada una unidad de acción.
- **FER:** Reconocimiento de expresiones faciales (por sus siglas en inglés *Facial Expresion Recognition*).
- **FPP:** Posición de puntos característicos, por sus siglas en inglés *Feature points positions*.
- **FPS:** Fotogramas por segundo (FPS, por sus siglas en inglés *frames per second*).
- **GDL:** Grados de libertad (por sus siglas en inglés *DOF, Degrees Of Freedom*) es el número de formas en que se puede mover un mecanismo.
- **GNU:** Licencia pública general, por sus siglas en inglés *General Public License*.
- **ICA:** Análisis de componentes independientes, por sus siglas en inglés *Independent Component Analysis*.
- **ICP:** Por sus siglas en inglés *Iterative closest point*, es un algoritmo empleado para minimizar la diferencia entre dos nubes de puntos.
- **IMT:** Por sus siglas en inglés *Interactive Motion Technologies*, es una empresa de robótica, pionera a nivel mundial.
- **LDA:** Análisis discriminante lineal, por sus siglas en inglés *Linear Discriminant Analysis*.
- **LDP:** Patrón direccional local, por sus siglas en inglés *Local Directional Pattern*.
- **LGC:** Código local de gradiente, por sus siglas en inglés *Local Gradient Code*.

- **NN:** Vecino más cercano, por sus siglas en inglés *Nearest Neighbor*.
- **NUI:** Interfaz de usuario natural, por sus siglas en inglés *Natural User Interface*, aplicación proporcionada por Microsoft para el reconocimiento de rostros.
- **PCA:** Análisis de componentes principales (por sus siglas en inglés *Principal Component Analysis*), es una técnica utilizada para reducir la dimensión de un conjunto de datos.
- **PHR:** Registro personal de salud, por sus siglas en inglés *Personal Health Register*.
- **Plantarflexión:** Acción de la articulación del tobillo en la que el pie se dirige hacia abajo, en un rango máximo de -45° .
- **RBF:** Kernel Gaussiano en las máquinas de soporte vectorial, por sus siglas en inglés *Radial Basis Function*.
- **ROI:** Regiones de interés, por sus siglas en inglés *Regions Of Interest*.
- **RSM:** Repetición de movimientos específicos, por sus siglas en inglés *Repetition of Specific Movements*.
- **SCRIPT:** Cuidado supervisado y rehabilitación involucrando a la telerobótica, por sus siglas en inglés *Supervised Care and Rehabilitation Involving Personal Telerobotics*.
- **SDK:** Kit de desarrollo de software, por sus siglas en inglés *Software Development Kit*.
- **Sensor RGB:** Es aquel que tiene la capacidad de visualizar imágenes conteniendo 3 colores básicos rojo, verde y azul. (Por sus siglas en inglés, Red, Green, Blue).
- **SG:** Por sus siglas en inglés, *Serious Games*, los juegos serios son sistemas interactivos, los cuales su objetivo principal es brindar alguna habilidad ya sea nueva o perdida. Se consideran multidisciplinarios ya que son desarrollados para distintas áreas como la rehabilitación o la educación.
- **SVM:** Máquinas de soporte vectorial, por sus siglas en inglés, *Support Vector Machines*, son un conjunto de algoritmos de aprendizaje supervisado.
- **Unidades de acción (UA):** Unidades de acción que van asociadas a diferentes movimientos y elementos del rostro, diseñadas por Paul Ekman.
- **XML:** Lenguaje de Marcado Extensible, por sus siglas en inglés *eXtensible Markup Language*.
- **XNA:** es un conjunto de herramientas, con un entorno de ejecución, administrado por Microsoft, que facilita el desarrollo y gestión de juegos de ordenador.

Capítulo 1: Introducción

En esta sección se describe la problemática de esta tesis, así como sus objetivos, alcances, limitaciones y justificación.

1.1 Problemática

La articulación del tobillo es una de las estructuras de mayor soporte del cuerpo humano, como resultado de su función es la articulación comúnmente más lesionada [1], el esguince de tobillo surge del desplazamiento hacia afuera del pie, rompiendo los ligamentos. Este problema puede ser clasificado en tres grados [2]:

- Grado 1: Estiramiento de los ligamentos.
- Grado 2: Algunas rupturas parciales en uno o más ligamentos.
- Grado 3: Ruptura completa de uno o más ligamentos.

La rehabilitación es el proceso en el cual a través de la fisioterapia es posible recuperar las habilidades físicas perdidas en un incidente traumático. Dependiendo del grado de la lesión serán las sesiones que se deben realizar, las cuales van entre las 5 y las 25 sesiones terapéuticas aproximadamente, si se lleva una adecuada rehabilitación, el tobillo podrá recuperar la fuerza y movilidad perdida sin convertirse en un problema crónico.

El tobillo tiene tres posibles movimientos, entre ellos se encuentra el movimiento de dorsiflexión-plantarflexión. Como se muestra en la Figura 1.1 durante la dorsiflexión el pie es atraído hacia arriba en dirección a la pierna mientras que en la plantarflexión en dirección opuesta. El intervalo de valor máximo para este movimiento es de -45° a 25° . Este tipo de ejercicio de rehabilitación es el más utilizado por el personal médico.

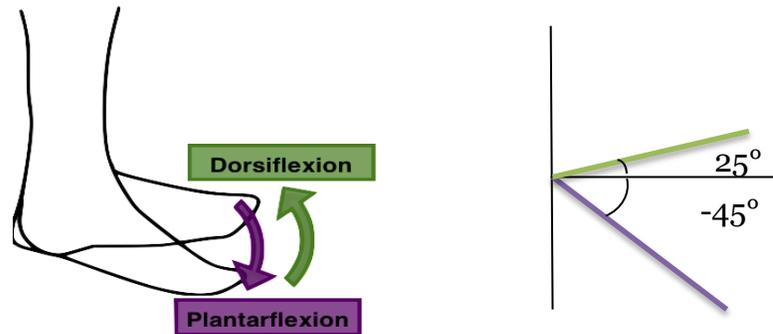


Figura 1.1 Movimiento de dorsiflexión-plantarflexión.

Existen cuatro niveles de recuperación dependiendo del nivel que el paciente requiera: pasiva, asistida, activa y resistiva [1].

La recuperación pasiva es aquella en la que el paciente no interviene en lo absoluto en la ejecución del movimiento, por lo general el terapeuta realiza todo el esfuerzo. Una modalidad intermedia es la asistida donde combina el esfuerzo del paciente y el terapeuta. La recuperación activa es cuando el paciente realiza todo el esfuerzo por sí mismo y en la resistiva se presenta una fuerza opuesta al movimiento del paciente.

Las principales lesiones que afectan a la región del tobillo son los esguinces y las fracturas. Ambas se producen principalmente en actividades deportivas; sin embargo, pueden también presentarse en la vida cotidiana ya que alguno de los ligamentos que sustentan la articulación sobrepasa sus límites de elasticidad, debido generalmente a un movimiento brusco y forzado.

Una lesión en tobillo puede aumentar el riesgo de una nueva lesión o lesiones repetitivas entre un 40% y un 70% [1]. Por esta razón es importante trabajar el fortalecimiento y el estiramiento de las estructuras cercanas a la articulación del tobillo, para disminuir el riesgo de una nueva lesión.

Por otro lado, los pacientes consumen tiempo, dinero y esfuerzo en las rehabilitaciones, ya que el equipo de rehabilitación consta de varios especialistas, haciéndolo un proceso muy costoso.

El problema que se presenta es el abandono del proceso de rehabilitación por parte de los pacientes con lesiones de tobillo, esto debido a la desmotivación que sufren

por el largo y tedioso proceso de rehabilitación y las molestias propias de las terapias, impidiendo recuperarse completamente.

Por lo que se considera necesario contar con un sistema integral de interacción que motive al paciente a llevar a cabo su rehabilitación de tobillo de manera constante, mediante un sistema integral que incluye un sistema de visión artificial que mediante la detección de expresiones faciales es capaz de detectar de manera automática si el movimiento de tobillo solicitado es agradable o molesto, la utilización de un juego serio cuya dificultad varía de acuerdo con lo estipulado por el terapeuta y con la información visual obtenida.

La complejidad fue reflejada en los siguientes problemas los cuales fueron resueltos para la realización de esta tesis:

- Investigar y proponer el reconocimiento de tres expresiones faciales diferentes a las básicas, a partir de la descripción de las acciones faciales de Ekman.
- Describir, visualizar y reconocer en tiempo real las expresiones faciales del paciente, ya que ésta información ayuda a delimitar la dificultad en que debe estar situado el juego serio y con ello el esfuerzo requerido en el rehabilitador.
- Elección y estudio de las herramientas determinadas para el desarrollo e implementación de los dos sistemas requeridos: el sistema de reconocimiento de expresiones faciales (FER) y el juego serio.
- Implementación del juego serio para la rehabilitación, desarrollando un juego atractivo con tres diferentes niveles de juego, pero también sencillo y amigable, capaz de tener una respuesta en tiempo real, capaz de retroalimentarse del reconocimiento de expresiones faciales del paciente.
- Integrar estos dos sistemas desarrollados con el sistema rehabilitador del tobillo de un grado de libertad.

1.2 Objetivos

1.2.1 Objetivo general

Desarrollar un sistema inteligente de interacción basado en un juego serio y un sistema de visión artificial para la detección de expresiones faciales del paciente, que mejore el proceso de rehabilitación de tobillo, contemplando el movimiento de dorsiflexión-plantarflexión.

1.2.2 Objetivos específicos:

- Estudiar el rehabilitador de 1 gdl.
- Revisar el estado del arte en las temáticas de reconocimiento de expresiones faciales y juegos serios para rehabilitación.
- Analizar las expresiones faciales considerando las unidades de acción.

- Estudiar el *sdk* de Kinect para el reconocimiento de expresiones faciales.
- Seleccionar el software de desarrollo para la implementación de los sistemas.
- Diseñar e implementar el juego serio para rehabilitación.
- Diseñar e implementar un sistema de visión artificial que identifique las expresiones faciales especificadas.

1.3 Alcances y limitaciones

1.3.1 Alcances

- Se detectan 3 expresiones faciales motivado, desmotivado y con dolor.
- El juego cuenta con 3 niveles, de los cuales se dividirá los grados de movimiento de dorsiflexión-plantarflexión en 3.
- El juego tiene una interfaz sencilla y amigable.
- Se considera las expresiones para los niveles del juego.
- Se realiza pruebas en simulación.
- Se trabaja con la información de Kinect.

1.3.2 Limitaciones

- Sólo se aborda el movimiento de dorsiflexión-plantarflexión correspondiente a 1 grado de libertad.
- El sistema sólo será probado en personas sanas.
- No se hace el reconocimiento de rostros.
- No se garantiza el correcto reconocimiento de expresiones faciales si las personas están usando accesorios (gorro, lentes, barba, bigote, etc.).
- Debe existir un solo rostro en cada secuencia de video.
- Se utiliza al menos 6 UA's que proporciona Kinect.
- No se aborda las técnicas de control del rehabilitador.

1.4 Justificación

Es necesario un proceso interactivo de rehabilitación para que el paciente que ha sufrido un incidente traumático tenga una recuperación completa, los juegos serios motivan al paciente lesionado a realizar el proceso de rehabilitación mediante el movimiento de dorsiflexión-plantarflexión, ayudando a recuperar la habilidad perdida del tobillo.

El uso de juegos serios aplicados es cada vez más usado, causando un gran impacto ante la sociedad, en la rehabilitación se puede observar que los pacientes son más constantes, ya que estos se pueden utilizar tanto en la clínica como en el hogar.

Además, es adaptable a lo que el paciente está sintiendo en tiempo real ya que el juego serio es retroalimentado con sus expresiones faciales mediante un sistema de

visión artificial utilizando Kinect, tales expresiones son: dolor, motivado y desmotivado.

Si el paciente se nota adolorido la frecuencia en que pasa un obstáculo con respecto a otro es mayor a la neutral, para que ya no tenga tanto dolor y pueda adaptarse al juego y si el paciente se encuentra distraído el obstáculo pasara en una menor frecuencia.

1.5 Organización de la tesis

Este trabajo está organizado de la siguiente forma:

- Capítulo 2: Contiene el estado del arte relacionado al reconocimiento de expresiones faciales, así como los algoritmos utilizados para el análisis de las dos diferentes bases de datos creadas y los esquemas de funcionamiento del sistema de reconocimiento de expresiones faciales.
- Capítulo 3: Este capítulo incluye la información acerca del estado del arte referente a juegos serios de rehabilitación, así como el diseño y desarrollo de los tres diferentes niveles basados en las etapas de la rehabilitación.
- Capítulo 4: En este capítulo se muestra el diseño y desarrollo de la integración de los tres sistemas (sistema de reconocimiento de expresiones faciales, juego serio para rehabilitación y rehabilitador de un grado de libertad).
- Capítulo 5: Describe las pruebas que se realizaron para mostrar la funcionalidad y eficiencia de los sistemas desarrollados. También se presenta el análisis de los resultados obtenidos.
- Capítulo 6: Contiene las conclusiones generales, aportaciones y los trabajos futuros que pueden realizarse a partir de esta tesis.

Capítulo 2: Reconocimiento automático de expresiones faciales

Las expresiones faciales en el ser humano son un mecanismo natural para mostrar satisfacción, disgusto, duda, etc., pues son medios de expresión que cualquier persona puede detectar sin importar raza, edad o sexo; sin embargo, esta tarea es aún compleja para los sistemas actuales de reconocimiento automático a través de unidades de acción o puntos específicos ya que determinar el estado interno de una persona mediante sus movimientos musculares faciales requiere ponderar muchísimas variables [5].

Existen varios factores que influyen en la formación de las expresiones faciales [6], los cuales se muestran en la Figura 2.1. Para este trabajo se utilizaron factores como estados mentales (para la motivación y distracción) y actividades fisiológicas como el dolor.

Las expresiones faciales pueden ser visualizadas mediante unidades de acción (UAs), las cuales son un sistema de codificación creado por Paul Ekman, que identifican los movimientos del rostro.

El reconocimiento de expresiones faciales FER (por sus siglas en inglés, *Facial Expression Recognition*) ha sido una línea de investigación abordada en las últimas

décadas, obteniendo información no verbal acerca del comportamiento de las personas, actualmente se han realizado sistemas de visión artificial que realizan una clasificación de las expresiones faciales mediante el modelado de las unidades de acción con diferentes algoritmos de clasificación.

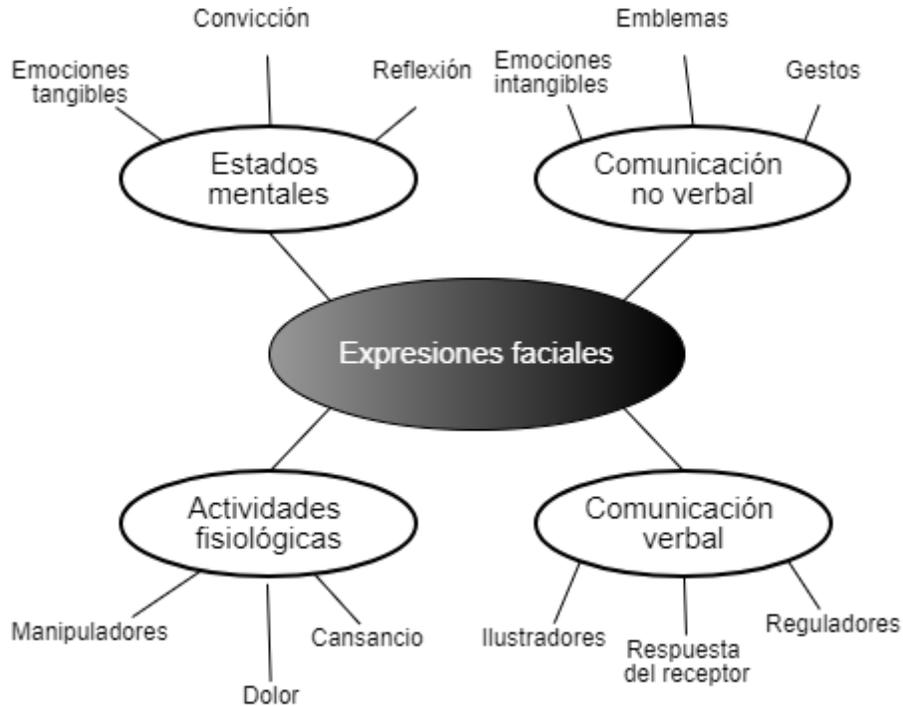


Figura 2.1 Factores que influyen en las expresiones faciales [6].

En este capítulo se presentan los trabajos que se consideran más importantes referentes al reconocimiento de expresiones faciales, también se muestra cómo se utilizó la librería *facetracking* de Kinect para la detección de los componentes del rostro y sus movimientos (UA), el desarrollo del sistema de reconocimiento de expresiones faciales, así como las expresiones faciales propuestas para su identificación en esta tesis.

2.1 Estado del arte referente a FER

Real time facial expressions recognition in video using support vector machines[7]

En este trabajo se realizó reconocimiento y clasificación de expresiones faciales mediante máquinas de soporte vectorial (SVM por sus siglas en inglés support vector machines, reconociendo las seis expresiones básicas (ira, asco, miedo, alegría, dolor y sorpresa).

Para su metodología utilizaron las etapas que se ven en la Figura 2.2:

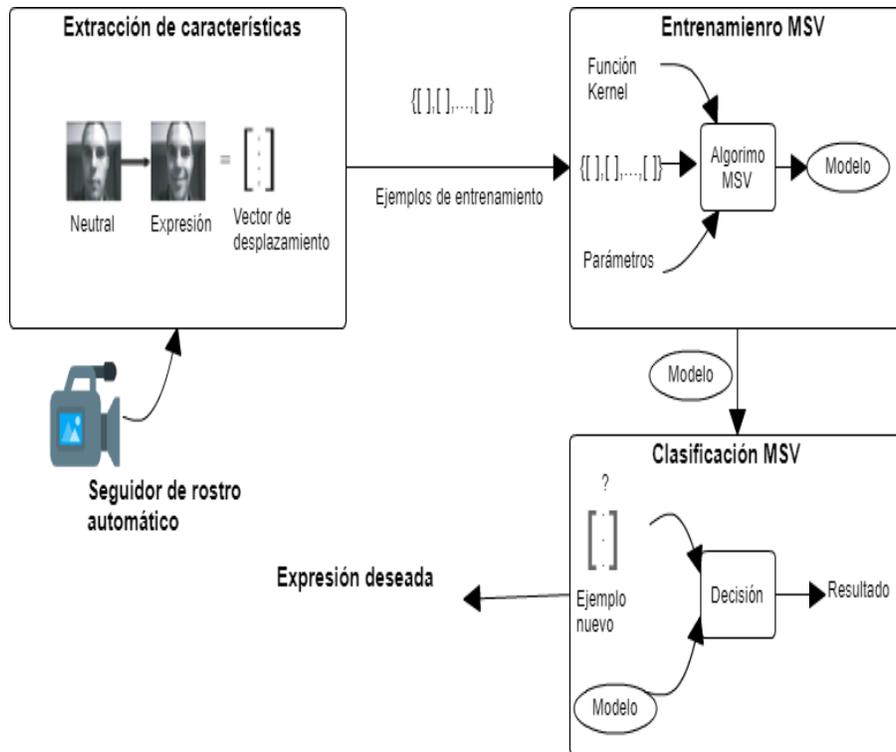


Figura 2.2 Metodología MSV[7].

La base de datos utilizada en este trabajo fue Cohn-Kanade.

Como conclusión en este trabajo se reconoce correctamente el 87.9% de las expresiones realizadas, teniendo un mayor problema con la expresión de temor con un porcentaje de reconocimiento de 71.7%.

Sistema de reconocimiento de expresiones faciales aplicado a la interacción humano-computadora usando redes neuronales y flujo óptico [5]

En este trabajo se implementó un sistema capaz de reconocer distintas expresiones faciales, basándose en la caracterización de las expresiones y del movimiento de la cara, a través de sucesiones de imágenes.

El autor propone investigar los movimientos faciales del ser humano y determinar el campo de acción de los mismos (unidades de acción) utilizando redes neuronales, siguiendo una metodología en la que se identifican las unidades de acción para posteriormente recopilar las características y entrenar el sistema para su reconocimiento. Para satisfacer este modelo de optimización, el autor definió cinco módulos principales:

- Examen facial: en este módulo se realiza un cuestionario con preguntas extremas al usuario, las cuales causarán alguna expresión facial que será capturada para el aprendizaje.
- Captura de caras: consiste en crear expresiones para el entrenamiento, se almacenan las imágenes en un formato en el que las pérdidas de información sean mínimas para clasificar los rasgos, posteriormente.
- Clasificación: su finalidad es disminuir el número de muestras que se requieren para entrenar la red neuronal, las características que clasifica el autor son los ojos y la boca.
- Entrenamiento: se toma como entrada el archivo de salida de la clasificación para entrenar la red neuronal.
- Reconocimiento: el propósito es utilizar la red neuronal ya entrenada para poder detectar en tiempo real una de las siguientes expresiones: distracción, duda, sorpresa, interés y pensando.

Para su desarrollo se utilizó *Visual Studio .NET* con la biblioteca de *OpenCV*.

El autor probó que el reconocimiento de expresiones faciales en tiempo real con redes neuronales es efectivo, con un 97% en el entrenamiento y para los patrones de verificación de al menos 75%.

Real time 2D-3D facial action and expression recognition [8]

En este trabajo se presenta un sistema de reconocimiento de unidades de acción y expresiones faciales para imágenes 2D y 3D, es un sistema basado en reglas para seguimiento del rostro y localización de características en tiempo real.

Para la codificación de los movimientos faciales se usó el sistema de codificación de acciones faciales FACS (*Facial Action Coding System*), el cual fue desarrollado por Ekman y Friesen, donde los movimientos del rostro son descritos en 44 unidades de acción.

El sistema está basado en reglas por lo que mediante ellas se compara si hay algún movimiento, de las unidades de acción respecto a una expresión neutral, y después se clasifica la unidad de acción o expresión facial que se está presentado.

Las pruebas fueron realizadas con una base de datos que consiste en 832 secuencias con 52 participantes (12 mujeres y 40 hombres) de entre 24 y 40 años, realizando expresiones de felicidad, tristeza, disgusto, sorpresa y neutral, teniendo un 83.6% de asertividad.

Facial expression recognition using facial characteristic points and Gini index [9]

Se presenta un sistema de reconocimiento de expresiones faciales utilizando características faciales y Gini index, utilizando ciertas regiones de la cara como lo son: las cejas, los ojos y la boca, ya que estas proporcionan una mayor información.

Para su entrenamiento se utilizó la base de datos *JAFFE* [10], la cual consta de 30 imágenes con las expresiones básicas (enojo, alegría, sorpresa, miedo, tristeza y neutral). La herramienta utilizada para el desarrollo de este sistema fue *Matlab*.

Los autores utilizan un árbol de decisión a partir del cual es elegida la expresión correspondiente a la imagen analizada, esto mediante los siguientes pasos: se mide la altura de los ojos, si están en un índice mayor a lo normal se dice que es una expresión de sorpresa. Si la respuesta es negativa, se ve si están abiertos; De acuerdo al porcentaje de apertura de los ojos, se asume que es una expresión de miedo y si están más cerrados o cerrados, se procede a ver su amplitud; si están en un rango de amplitud más que lo normal (tiende a ser más rasgados) se dice que está en una expresión de felicidad y se mide la amplitud de la boca. Si la boca tiene una amplitud mayor a un umbral, se asume que esta triste, pero si esta menor se pasa a ver si la boca se encuentra abierta donde si es muy poca la abertura se considera que la persona está enojada, y si no, se considera que se encuentra en un estado neutral.

ROI segmentation for feature extraction from human facial images [11]

El presente trabajo presenta una técnica de extracción de características de regiones de interés del rostro (ROI, por sus siglas en inglés *Regions Of Interest*) con ayuda de una segmentación por el color de la piel.

La metodología propuesta por los autores se describe en las siguientes etapas:

- Leer una imagen.
- Aplicar segmentación por el color de la piel.
- Detección de las regiones interconectadas.
- Detección del rostro en la imagen.
- Crear una imagen binaria a partir de la RGB.
- Detección del rostro en la imagen binaria.
- ROI.
- Identificación de características.

Los métodos basados en características tienen eficientes resultados, utilizando este método se pueden encontrar puntos específicos en la región entera de acuerdo con

el área de interés. Estos puntos son llamados *feature points* (puntos característicos), para obtenerlos es necesario la detección del rostro en la imagen y luego por regiones específicas (ojos y labios) encontrar dichos puntos característicos.

Sistema de reconocimiento de rostros en 3D usando Kinect [12]

En este trabajo se implementa un sistema de reconocimiento facial basado en características 3D, para el cual utilizan Kinect, proporcionando imágenes a través de la *API* ofrecida por Microsoft.

Para su metodología se utilizó eigenfaces y PCA (análisis de componentes principales).

El desarrollo de este sistema se llevó a cabo en la *API NUI (application programming interface, natural user interface)* proporcionada por Microsoft, también fueron utilizadas librerías como MEX, ya que es necesario el crear un puente entre el lenguaje C que tienen las capturas con Kinect y *Matlab* para su análisis.

Se creó la base de datos en dos secciones, una para cabello largo y la otra para cabello corto, tomando en cuenta las siguientes consideraciones:

Cabello largo: 2 fotos con cabello suelto, sin gafas y sin sombrero; 3 con cabello suelto, con gafas y sin sombrero; 3 con cabello suelto, con gafas y con sombrero; 3 con cabello recogido, sin gafas y sin sombrero; 3 con cabello recogido, con gafas y sin sombrero; 1 con cabello recogido, con gafas y con sombrero.

Cabello corto: 3 fotos sin gorra, sin gafas y sin sombrero; 3 sin gorra, con gafas y sin sombrero; 3 sin gorra, con gafas y con sombrero; 3 con gorra, sin gafas y sin sombrero; 3 con gorra, con gafas y sin sombrero.

Los resultados indican que el uso de imágenes a profundidad aporta mayor información útil que puede mejorar el desempeño de técnicas de reconocimiento facial como PCA y eigenfaces.

Unsupervised facial expression recognition and avatar reconstruction from Kinect [13]

En este artículo se hace la detección de seis expresiones faciales (sorpresa, temor, disgusto, enojo, tristeza y alegría) automáticamente, las cuales son reconstruidas en un avatar.

También crearon su propia base de datos con 20 participantes, la cual contiene imágenes y nubes de puntos de localización de expresiones faciales.

Su metodología se basa en identificar el rostro y segmentarlo por regiones, para la detección facial se usan clasificadores ya implementados en *openCV* y para la detección de características lo que se busca primordialmente son los ojos y la boca por lo que se usan clasificadores en cascada (*haarcascade*).

Para el reconocimiento de la expresión facial se utiliza el método eigenfaces.

Real-time performance-driven facial animation with 3ds Max and Kinect [14]

En este trabajo se encuentra como problemática el mapeo de expresiones faciales de un actor para que sean representadas por un avatar, por lo que su objetivo principal es brindar una animación en tiempo real del rostro usando *3ds Max* y *Kinect*.

La metodología propuesta consta de 4 pasos:

1. Captura de video en 3D: se toma un video del usuario usando Microsoft Kinect, usando simultáneamente captura en 2D a color y en mapa de profundidad (30 cuadros por segundo).
2. Facetracking: se predefine un grupo de puntos en el rostro; computando la posición de la cabeza y las expresiones, para lo cual utiliza AAM (*Active Appearance Model*) que es un algoritmo de visión artificial popular para el *facetracking* y ICP (*Iterative closest point*) para calcular las unidades de acción con los puntos en 2D.
3. Transmisión de datos: se construye el canal de comunicación para la transmisión del control de parámetros de *facetracking* al software *3Ds Max* para la animación del rostro en tiempo real.
4. Síntesis de la expresión facial: los parámetros de control son útiles para usar el modelo facial en el avatar, las expresiones que hagan las reproducirá el avatar en tiempo real.

Las herramientas de desarrollo de este trabajo son: *visual studio 2010*, *Kinect SDK*, *3Ds Max*.

Facial expression recognition using Kinect depth sensor and convolutional neural networks [15]

En este trabajo se implementó un sistema de reconocimiento facial utilizando redes neuronales convolucionales (CNN) con un sensor Kinect.

Proponen la arquitectura donde primero se adquiere una imagen facial a través de Kinect, se preprocesa, se hace la representación de la imagen, se clasifica en la red neuronal convolucional y luego se detecta la expresión facial.

Los autores hacen notar que a pesar de que las imágenes que utilizaron (*EURECOM Kinect face dataset*) tienen ruido, el reconocimiento de expresiones faciales a través de su red neuronal convolucional es muy acertado. Para su desarrollo se utilizó *Kinect SDK*.

3D facial expressions recognition using the Microsoft Kinect [16]

En este trabajo se analizó la aplicación de reconocimiento de expresiones faciales desde dos métodos diferentes:

- FACS (Sistema de codificación de acción facial, el cual su reconocimiento es en tiempo real.
- Momentos de Zernick Local (MZ's), reconocimiento en tiempo no real.

Para el reconocimiento en tiempo real el autor usó la librería *facetracking* proporcionada por Kinect que ofrece la detección de 6 UAs, las cuales ayudan a identificar las expresiones faciales básicas, sin embargo, para que haga reconocimiento de las mismas se tienen que realizar movimientos exagerados, teniendo por lo mismo resultados limitados

Para el método que no es en tiempo real, realizó un preprocesado localizando la nariz y poniendo una esfera sobre ella para recortar su imagen en 3D a 2D. Luego procesa la imagen con los momentos de Zernick, localiza y extrae características (partes del rostro) para después compararlas con una plantilla y clasificar.

Los experimentos se realizaron con las bases de datos *VBN* [17] y *FaceWarehouse* RGB-D [18] con las cuales se tuvo un 87.68% de reconocimiento.

Adaptive feature selection and data pruning for 3D facial expression recognition using Kinect [19]

El interés de este trabajo es el desarrollar un sistema que reconozca las distintas expresiones faciales principales (felicidad, tristeza, miedo, enojo y desagrado) usando un sensor 3D Kinect.

El método que propone el trabajo se lleva a cabo en tres etapas:

- 1) *Registro de datos*: se usa el método de Procustes de superposición ya que es muy eficiente para la nube de puntos; Procustes encuentra una transformada lineal que mejor se ajuste a un conjunto de puntos de referencia con una distancia euclidiana.
- 2) *Extracción de características*: esta etapa está dividida en tres subetapas:
 - I. Triangulación de Delanuy sobre los puntos clave en 3D, para obtener una malla fija (T) de triángulos que conecta los puntos clave de la cara.
 - II. Se extrae un conjunto de aristas (E) de la malla triangular.

- III. Para cada arista en E se calcula su longitud usando 9 modalidades: coseno, estandarizado euclidiano, distancia Manhattan, Chebychev, Minkowski, Hamming, Spearman, Correlación y Jaccard.
- 3) *Clasificación*: para la clasificación de expresiones se utilizan las máquinas de soporte vectorial.

Para la experimentación se utilizaron dos conjuntos de datos (*datasets*), el primero es FaceWarehouse [18] y el segundo fue un *dataset* propio, el cual contiene 17,000 imágenes 3D de 10 sujetos entre frontales y no frontales.

Se realizó experimentación con poda de datos solo en las imágenes de frente, utilizando solo 20% de los puntos dados, obteniendo más del 96% de precisión en la clasificación.

Después del entrenamiento y pruebas con ambos *datasets* se pudo observar que se tiene una precisión del 95.1% en posiciones no frontales y un 98% con posiciones frontales.

A real time facial emotion recognition using depth sensor and a interacting with second life based virtual 3D avatar [20]

Los autores de este trabajo detectaron las emociones básicas mediante la librería *Facetracking* de *Kinect SDK*, la cual utiliza la máscara Candide-3 para obtener puntos verticales como lo son las cejas y la boca. También utilizaron el sistema FACS para reconocer las expresiones de felicidad, miedo, ira, tristeza y sorpresa.

A partir de estas se creó un avatar para una aplicación llamada *second life* (un mundo virtual en el que se puede interactuar con otras personas), teniendo como fin el representar las emociones que se están sintiendo en ese momento.

Los problemas que presenta este trabajo es que no puede reconocer la expresión de disgusto, tomándola como tristeza u enojo; además, para que se puedan reconocer las expresiones del individuo se debe de tener al menos un metro de distancia con el sensor Kinect.

VT-KFER: A Kinect-based RGBD+Time Dataset for Spontaneous and Non-Spontaneous Facial Expression Recognition [21]

En este trabajo se crea una base de datos llamada VT-KFER la cual contiene las seis expresiones faciales básicas capturadas con la cámara RGB de Kinect:

- Felicidad.
- Enojo
- Temor.
- Disgusto.

- Tristeza.
- Sorpresa.

VT-KFER incluye 1956 secuencias de imágenes RGB y mapas de profundidad para seis expresiones en tres posiciones diferentes, además de tres intensidades de luz.

Esta base de datos fue creada con 32 participantes con diferentes características, como el género, la nacionalidad, si tenían alguna oclusión facial (bigote, barba, lentes) y edades, también se menciona que es la primera en incluir niños.

Para su creación se utilizaron dos tipos de expresiones:

- Expresiones espontáneas: mostraron diferentes imágenes a los integrantes y estas debían causar cierta emoción en ellos.
- Expresiones no espontáneas: se pidió a los participantes que hicieran seis expresiones en tres diferentes posiciones (frontal, perfil derecho y perfil izquierdo).

Para su desarrollo se utilizó Kinect, capturando imágenes en RGB-D.

Using Kinect for real-time emotion recognition via facial expression [22]

En este artículo se propone capturar imágenes en 3D fusionando las unidades de acción con FPPs (por sus siglas en inglés *Feature point positions*, posición de puntos característicos) proporcionados por Kinect (alzar cejas, bajar cejas, alzar labios, estirar labios, hundir labios y bajar barbilla).

Con este sistema se pueden detectar las 6 emociones básicas (enojo, disgusto, temor, felicidad, tristeza y sorpresa).

La arquitectura utilizada en su sistema primero adquiere una secuencia de video con el Kinect, en la que se detecta la cara y se empieza a hacer la extracción de características tanto con las unidades de acción como con FPP, posteriormente realizan un análisis de esta información y se hace una estimación final tomando en cuenta ambos resultados parciales; para la clasificación se usa una máquina de soporte vectorial.

Al fusionar estos dos algoritmos se logró tener una detección de expresiones faciales en tiempo real.

Facial expression recognition: A survey [23]

Uno de los métodos no verbales de comunicación para entender el estado de ánimo y mental de una persona son las expresiones faciales, el reconocimiento de

expresiones faciales FER (por sus siglas en inglés *facial expression recognition*) se ha convertido en una interesante área de visión artificial y sus áreas de aplicación no están limitadas al estado de ánimo de las personas, seguridad, sistemas de orientación automática, detección de mentiras, etc. FER está constituido por 5 etapas que se muestran en la Figura 2.3:



Figura 2.3 Etapas FER.

La detección de componentes faciales se realiza a través de las regiones de interés del rostro (*ROI*) las cuales son los ojos, la boca, nariz, cejas, etc.

Las técnicas más populares para la extracción de estas características son: los filtros Gabor, patrones binarios locales, análisis de componentes principales (*PCA*, *principal component analysis*), análisis de componente independiente (*ICA*, *independent component analysis*), análisis discriminante lineal (*LDA*, *linear discriminant analysis*), código local de gradiente (*LGC*, *local gradient code*) y patrón direccional local (*LDP*, *local directional pattern*).

Algunos métodos utilizados para la clasificación son: máquinas de soporte vectorial (*SVM*, *support vector machine*) y vecino más cercano (*NN*, *nearest neighbor*).

Las expresiones faciales están categorizadas en seis emociones con 13 unidades de acción, las cuales se pueden observar en la Tabla 2.1.

Tabla 2.1 Expresiones faciales básicas.

Emoción	Combinación de unidades de acción
Enojo	Frente abajo, parpado superior arriba, parpado tenso, labios tensos.
Disgusto	Nariz arrugada esquina del labio abajo, labio inferior abajo.
Temor	Ceja interior/exterior levantada, frente abajo parpado superior tenso, parpado apretado, labios estirados, mandíbula caída.
Felicidad	Mejilla levantada, esquina del labio estirada.
Tristeza	Frente media levantada, lados de la frente abajo, labios hacia abajo.
Sorpresa	Frente levantada, parpado superior levantado, mandíbula caída

En la Tabla 2.2 se muestran los resultados de la experimentación de este trabajo, donde se probaron varios métodos para el reconocimiento, utilizando la base de datos *JAFFE* [10] la cual tiene 213 imágenes de 10 objetos con las 6 expresiones básicas y la neutral. Después de la detección del rostro usan el algoritmo *Viola Jones*, donde observan que el mejor método para este trabajo fue el gradiente local con diagonal en vertical (Tabla 2.2).

Tabla 2.2 Experimentación survey de FER.

Métodos	Rango de reconocimiento
LBP Local Binary Pattern	89.4231%
LGC Local Gradient Code	90.3846%
LGC-HD Local Gradient Code-Horizontal Diagonal	87.50%
LGC-VD Local Gradient Code-Vertical Diagonal	92.3077%
HOG Histogram of Oriented Gradients	85.7143%
LDP Local Direction Pattern	85.2041%

2.2 Análisis del estado del arte referente a FER

Las expresiones faciales en el ser humano son un mecanismo natural para mostrar satisfacción, disgusto, duda, etc., pues son medios de expresión que cualquier persona puede detectar sin importar raza, edad o sexo. Sin embargo, esta tarea es compleja para los sistemas de reconocimiento de expresiones faciales existentes pues determinar el estado interno de una persona mediante sus expresiones requiere ponderar muchísimas variables [20].

En la Tabla 2.3 se muestra un resumen del estado del arte referente al reconocimiento de expresiones faciales en el cual se puede observar las principales técnicas utilizadas para las distintas etapas de los sistemas, así como las bases de datos públicas más usadas.

En recientes años se puede observar cómo ha aumentado el uso del sensor Kinect para el reconocimiento de dichas expresiones, esto mediante el software de desarrollo proporcionado por Microsoft, *Kinect sdk*, el cual a su vez proporciona la herramienta de desarrollo *Facetracking sdk* para el seguimiento del rostro mediante las cámaras de Kinect.

El sensor Kinect es utilizado para el seguimiento del rostro y la detección de sus componentes, posteriormente se extraen las características (Unidades de acción) y a partir de ellas se realiza una clasificación en las diferentes expresiones faciales, las cuales son en la mayoría de los artículos revisados las básicas (enojo, felicidad, tristeza, temor y disgusto).

Durante el desarrollo de este proyecto de tesis se utilizó el sensor Kinect, así como la librería *Facetracking sdk* para el reconocimiento de expresiones faciales, las cuales estarán agrupadas en tres clases: motivado, desmotivado y con dolor; el reconocer dichas expresiones ayudará como retroalimentación al juego serio, sugiriendo en qué nivel el paciente se debe de encontrar.

Tabla 2.3 Resumen del estado del arte FER.

Autor	Área	Base de datos	Técnica	Precisión	Herramientas	Tiempo real	Expresiones
Michel (2003) [7]	FER	Cohn-Kanade	MSV	87.9%	N/A	Sí	Enojo, disgusto, temor felicidad, dolor, sorpresa.
Porras (2005) [5]	FER	Propia	Redes neuronales	N/A	Visual Studio .Net, OpenCV	Sí	Distracción, duda, interés, pensativo, sorpresa
Tsalakanidou (2010) [8]	Reconocimiento de acciones faciales	Propia	ASM (Active Shape Model)	83.6%	N/A	Sí	Neutral, disgusto, felicidad, sorpresa
Gupta (2012)[9]	FER	JAFFE	Árbol de decisión	86.66%	Matlab	No	Neutral, enojo, temor, felicidad, tristeza sorpresa
Surbhi (2012)[11]	Extracción de características	N/A	ROI	N/A	N/A	No	N/A
Arenas (2012)[12]	Reconocimiento de rostros	Propia	PCA. eigenfaces	N/A	API NUI, Kinect SDK, Matlab	No	N/A
Seddik (2013) [13]	FER	JAFFE	PCA. eigenfaces	64%	OpenCV, Kinect SDK, Visual studio 2010	Sí	Enojo, disgusto, temor, felicidad, tristeza, sorpresa
Lí (2013)[14]	Mapeo de expresiones faciales	N/A	AAM (Active Apperence Model), ICP (Iterative Clossesst Port)	N/A	3ds Max, Kinect SDK, Visual studio 2010	Sí	N/A
Ijjina(2014) [15]	FER	EUROCOMM	Redes neuronales convolucionales	87.99%	Kinect SDK	N/A	Neutral, felicidad y con la boca abierta
Stocci (2014) [16]	FER	VBN, FaceWarehouse	Facetracking, Momentos de Zernick	87.66%	Kinect SDK	Uno de los dos metodos sí	Enojo, felicidad, tristeza, sorpresa
Aly (2014) [19]	FER	FaceWarehouse, propia	MSV	95.1% no frontales y 98% frontales	Kinect SDK	Sí	Enojo, disgusto, temor, felicidad, tristeza, sorpresa
Kakarla (2014) [20]	FER	Facetracking	GEMEP-FERA	90%	Kinect SDK	Sí	Enojo, disgusto, temor, felicidad, tristeza, sorpresa
Aly (2015) [21]	FER	Propia (VT-KFER)	N/A	N/A	Kinect SDK	N/A	Enojo, disgusto, temor, felicidad, tristeza, sorpresa
Mao (2015) [22]	FER	Propia	MSV	87%	Kinect SDK	Sí	Enojo, disgusto, temor, felicidad, tristeza, sorpresa
Kumary (2015)	FER	JAFFE	LBP, LGC, LGC-HD, LGC-VD, HOG-LDP	89.4231%, 90.3846%, 87.5%, 92.3%, 85.7143%, 86.2041%	N/A	N/A	Enojo, disgusto, temor, felicidad, tristeza, sorpresa

También se puede ver que para la clasificación y pruebas experimentales del reconocimiento de expresiones faciales se debe usar una base de datos, que proporcione la información para que el sistema aprenda y posteriormente evalúe qué tan acertado es el sistema; entre los trabajos relacionados se puede observar que una de las más populares es JAFFE (*Japanese Female Facial Expressions*), la cual, es una base de datos que incluye imágenes de solo mujeres.

En el Anexo 1 se muestra un resumen de bases de datos de expresiones faciales públicas encontradas, como se puede observar, las expresiones consideradas en este trabajo no se encuentran en dichas bases, por lo cual se creó una base de datos propia.

2.3 Sensor Kinect

El sensor Kinect (Figura 2.4) fue construido para revolucionar la manera de jugar videojuegos y la experiencia de entretenimiento, sin embargo, ya no es el único objetivo de Kinect. En los últimos años se ha utilizado para otras áreas como el reconocimiento de expresiones faciales, ayudando a obtener información no verbal de lo que está sintiendo una persona.



Figura 2.4 Sensor Kinect.

Para utilizar la librería Kinect *facetracking* es necesario tener los siguientes requerimientos mínimos:

- Computadora dual-core, 2.66 Ghz.
- Windows 7, 8 o posterior.
- GB en RAM.
- Sensor Kinect (usb y power cabling).

Paul Ekman diseñó un sistema de codificación teniendo como objetivo clasificar los principales movimientos del rostro a través de las unidades de acción, el cual contiene 46 unidades de acción distintas y 14 códigos de motores de cabeza [4].

Kinect a través de la librería *facetracking* permite el seguimiento de rostros mediante la máscara candid-3 en tiempo real, identificando 3 movimientos de cabeza y 6 diferentes unidades de acción las cuales en combinación determinan diferentes expresiones faciales.

La UA0 en Kinect *facetracking* es referida a levantar el labio superior, la UA1 a bajar la mandíbula, la UA2 a apretar los labios, la UA3 a bajar las cejas, la UA4 a bajar las comisuras de los labios y la UA5 a levantar las cejas, estas unidades de acción son medidas en un rango desde -1 hasta 1 (Figura 2.5).

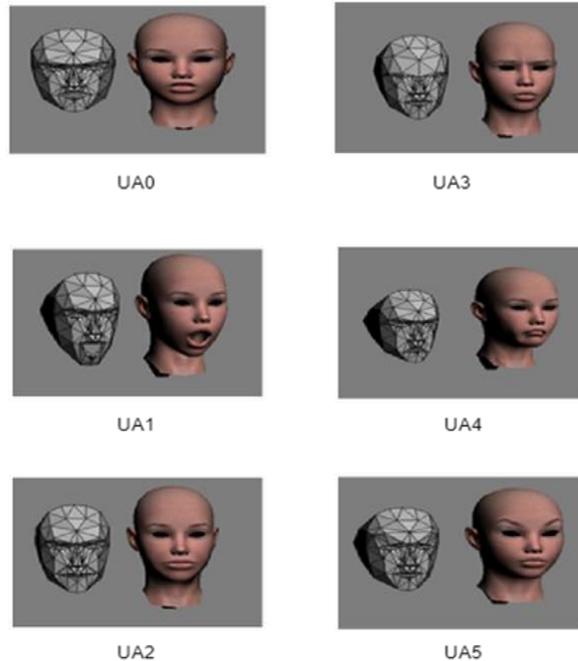


Figura 2.5 Unidades de acción identificadas con Kinect.

Los movimientos de cabeza que reconoce Kinect son medidos en un rango desde -90° hasta 90° . Teniendo tres posibles movimientos: *Pitch* el cual consiste en subir y bajar la cabeza, *Yaw* girar la cabeza hacia los costados y *Roll*, girar la cabeza (Figura 2.6).

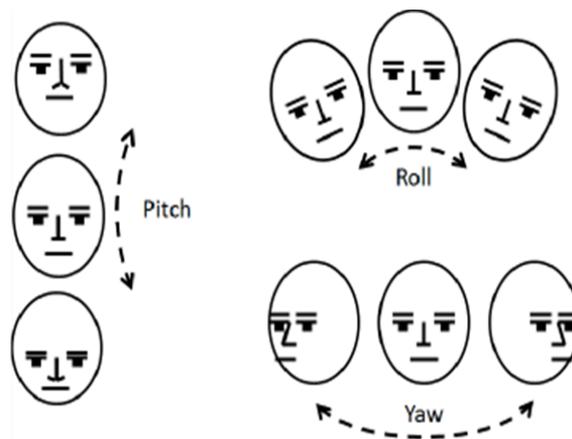


Figura 2.6 Movimientos de cabeza identificados con Kinect.

En la Tabla 2.4 se muestra una comparación entre las unidades de acción diseñadas por Paul Ekman y las proporcionadas por Kinect.

Tabla 2.4 Comparación de unidades de acción.

Kinect	Ekman	Descripción
UA0	UA5	Levantamiento de parpado superior
UA1	UA26	Caída de mandíbula
UA2	UA20	Apretar los labios
UA3	UA4	Bajar las cejas
UA4	UA15	Esquinas del labio abajo
UA5	UA2	Levantamiento exterior de la ceja

En la Tabla 2.5 se muestran las expresiones faciales básicas, las cuales son: enojo, disgusto, temor, felicidad y tristeza; y las unidades de acción que se requieren para reconocerlas en el sistema propuesto por Ekman y Kinect.

Tabla 2.5 Comparación de unidades de acción en las expresiones faciales básicas.

Emociones	Unidades de acción	UA's en Kinect	Ilustración
Enojo	UA4: bajar las cejas UA5: levantamiento del parpado superior UA7: apretar parpados UA23: apretar los labios	UA3: bajar las cejas	
Disgusto	UA9: arrugar la nariz UA15: esquina del labio abajo UA16: labio inferior abajo	UA4: Comisuras de los labios abajo	
Temor	UA1: levantamiento interior de la ceja UA2: levantamiento exterior de la ceja UA4: bajar cejas UA5: levantamiento del parpado superior UA7: apretar parpados UA20: apretar labios UA26: caída de mandíbula	UA2: apretar labios UA1: Caída de mandíbula	
Felicidad	UA6: levantamiento de mejillas UA12: estiramiento de esquina labial	UA4: esquina del labio abajo (valores aproximados a -1)	
Tristeza	UA1: levantamiento interior de la ceja UA4: bajar cejas UA15: esquina del labio abajo	UA3: bajar cejas UA4: esquina del labio abajo	

Para la detección de estos movimientos y unidades de acción se utiliza Candide-3, la cual es una máscara parametrizada de 113 vértices y 168 superficies (Figura 2.7).

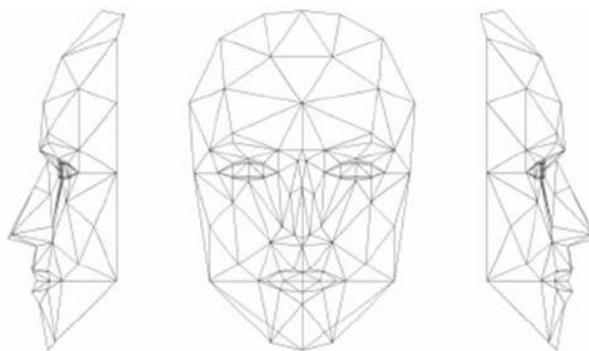


Figura 2.7 Candide-3.

2.4 Expresiones faciales propuestas

Para definir cuáles y cuantas expresiones faciales era necesario reconocer, se realizó un análisis de los requerimientos para retroalimentar el rehabilitador de tobillo y se consultó a un experto en psicología¹. Con la información recabada se propuso el reconocimiento de las siguientes tres expresiones faciales no básicas:

1. Dolor: esta expresión determinará si el paciente está sufriendo alguna molestia que le impida realizar cierto movimiento ya que se está forzando de más.
2. Motivación o concentración: en la que se puedan mostrar expresiones neutrales o de felicidad al estar jugando.
3. Desmotivado o distracción: en esta clase el paciente se encontraría distraído por lo que estaría volteando a los lados o de arriba abajo.

Para llevar a cabo la descripción de estas expresiones, se revisó y seleccionó un conjunto de las unidades de acción diseñadas por Ekman, que generalmente son utilizadas para describir las expresiones faciales básicas. El vector de características de cada expresión se integra de ocho atributos: 6 unidades de acción y 2 posiciones de cabeza, como se muestra en la Figura 2.8.



Figura 2.8 Vector de características.

En la Tabla 2.6 se muestra la descripción de las expresiones faciales propuestas para su reconocimiento y las unidades de acción utilizadas en cada una de ellas.

¹ Lic. Eduardo Sánchez Rivera, Licenciado en Psicología, Universidad del Valle de México.

Tabla 2.6 Descripción de las clases propuestas.

Clase	Descripción
Dolor (Clase 1)	<ul style="list-style-type: none"> • Bajar las esquinas de los labios (UA4) con valores mayores a 0.01. • Bajar las cejas (UA3) y dejar caer la mandíbula (UA1), ambas con valores mayores a 0.25. • Bajar cejas (UA3) y apretar los labios (UA2), ambas con valores mayores a 0.25.
Desmotivado o distracción (Clase 2)	<ul style="list-style-type: none"> • Volteando hacia arriba (Pitch): valores mayores a 35°. • Volteando hacia abajo (Pitch): valores mayores a -10°. • Volteando a la derecha (Yaw): valores mayores a 35°. • Volteando a la izquierda (Yaw): valores menores a -25°.
Motivado (Clase 3)	<ul style="list-style-type: none"> • Neutral: Todas las unidades de acción con valores que van de -0.20 a 0.20. • Feliz: Bajar las esquinas de los labios (UA3) con valores menores a -0.35.
Transiciones (Clase 4)	<ul style="list-style-type: none"> • Aquellas expresiones que no están incluidas en ninguna de las clases anteriores.

Como se observa en la Tabla 2.6 se agregó una categoría más. Esta clase es necesaria debido a que el rostro es un objeto variable que está en constante cambio de manera imperceptible como pestañear, movimiento de los ojos, cabeza, bostezar, ladear la cabeza, fruncir el ceño, etc. Entonces, para estas acciones normales, pero no contempladas, el sistema es capaz de dar una respuesta válida y no altera el sistema de reconocimiento de las expresiones faciales propuestas.

2.5 Esquema de funcionamiento del sistema de expresiones faciales

El funcionamiento del sistema de reconocimiento de expresiones faciales se realizó en dos etapas una denominada de entrenamiento y la segunda de validación. En el esquema mostrado en la Figura 2.9 se muestra la etapa de entrenamiento, en la cual se analizaron tres diferentes métodos de clasificación.



Figura 2.9 Esquema de funcionamiento para la etapa de entrenamiento del sistema de reconocimiento de expresiones faciales.

La adquisición de los videos de entrada se realiza utilizando el sensor Kinect en un ambiente sin controlar, posteriormente, se lleva a cabo la detección del rostro y sus componentes utilizando la máscara candide-3. La descripción de los mismos se realiza a través de las unidades de acción seleccionadas y los datos son almacenados en un archivo de texto, para después ser utilizados como entrada para llevar a cabo el entrenamiento con diferentes algoritmos de clasificación (C4.5, Maquinas de soporte vectorial y Naïve Bayes), creando modelos de clasificación para su uso en el sistema de reconocimiento de expresiones faciales.

Posteriormente, se llevó a cabo una segunda etapa: el objetivo de funcionamiento de este módulo fue evaluar los modelos de clasificación obtenidos en la etapa anterior (primer esquema). Para esta segunda etapa se creó una base de datos diferente a la considerada en la etapa de entrenamiento, esta base de imágenes se realizó con personas distintas y en un ambiente diferente (Ver Figura 2.10).

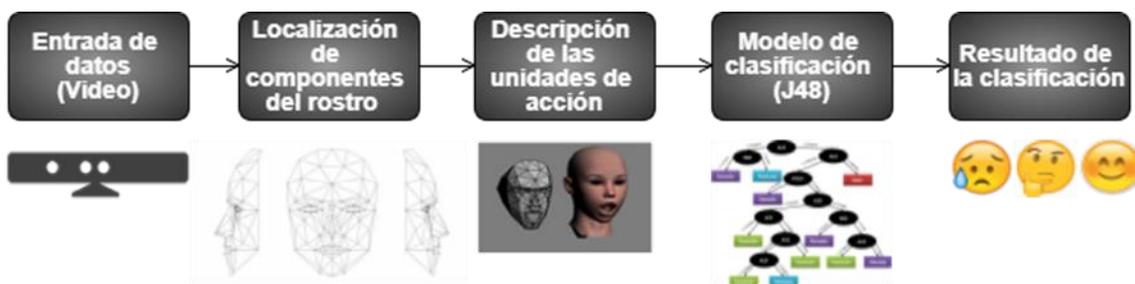


Figura 2.10 Esquema de funcionamiento para validar los modelos de clasificación para el sistema de reconocimiento de expresiones faciales (Etapa de prueba).

2.6 Creación de las bases de datos

Se llevó a cabo la creación de dos bases de datos, una para el entrenamiento y otra para la etapa de evaluación. La primera base de datos se integra de la captura de un total de 30 personas (10 mujeres y 20 hombres) entre los 22 Y 30 años. Se llevó a cabo en el laboratorio de inteligencia artificial (del cenidet) en un horario de 10 a 12

horas, en un ambiente sin controlar, pero con iluminación artificial (con las lámparas del laboratorio encendidas). Los videos se encuentran en formato RGB-D de 640X480 pixeles.

La adquisición se realizó a 30 FPS (Fotogramas por segundo). No obstante, los archivos de datos se integran del muestreo de la información de 7 imágenes por segundo, con un total de 306 instancias.

Analizando la información obtenida, se encontró que el rostro es un objeto muy variable y que está en constante cambio de manera imperceptible como pestañear, movimiento de los ojos, de cabeza, etc. Además de tener transiciones para pasar de una expresión a otra, por lo que para evitar el ruido y tener la expresión facial más marcada se consideró tomar para cada segundo solo una expresión facial. Entonces, con los 7 vectores de características almacenados por segundo, se realiza una votación y aquella expresión facial que tenga un valor más representativo con respecto a las clases propuestas, será la que asigne como clase al rostro para este segundo (ver Figura 2.11).



Figura 2.11 Almacenamiento.

En la Tabla 2.7 se muestra un ejemplo de cómo se realiza la selección de un vector, en el cual se considera un bloque de 7 vectores. Primeramente se revisa la clase y se selecciona aquella que es repetida un mayor número de veces, en este bloque la clase predominante es la clase 1, a partir de esto se eligió el vector que tuviera el valor más significativo para la clase 1, la UA4 que tiende a 1 significa que la persona esta triste (adolorido), es por eso que se eligió el vector número 6 ya que tiene la unidad de acción 4 igual a 0.3168, siendo el valor que más se acerca a 1.

Tabla 2.7 Ejemplo de selección.

UA0	UA1	UA2	UA3	UA4	UA5	PITCH	YAW	CLASE
-0.137457	0.0586584	-1	0.067049	0.26407	0.0450383	-1.1204	3.75883	Class_1
-0.137457	0.0586584	-1	0.067049	0.26407	0.0450383	-1.1204	3.75883	Class_1
-0.137457	0.0586584	-1	0.067049	0.26407	0.0450383	-1.1204	3.75883	Class_1
-0.20391	0.0909321	-1	0.0800984	0.301061	0.0491241	-0.92043	2.79804	Class_1
-0.247344	0.0835355	-1	0.0981119	0.294996	0.0784893	-0.85084	1.94085	Class_1
-0.340345	0.432133	-1	0.08894	0.31681	0.0838039	-0.4673	-0.165226	Class_1
-0.522762	0.0187934	-1	0.074292	0.205399	-0.0095483	0.42996	-9.10585	Class_1

Para la base de entrenamiento, a cada vector de características se le asignó una clase (*ground truth*) de acuerdo a las reglas inicialmente propuestas y que se pueden observar en la Tabla 2. Entonces, con los 7 vectores de características almacenados por segundo, se realiza una votación y aquella expresión facial que tenga un valor más representativo con respecto a las clases propuestas, será la que asigne como clase, al rostro para este segundo. El archivo de datos se integra de 2303 instancias.

La base de pruebas se integra de la adquisición de 15 videos de 15 personas distintas a las consideradas en la base de entrenamiento. La adquisición se realizó en el laboratorio de software (del Cenidet), en un horario de 14 a 16 horas, en un ambiente sin controlar, pero con iluminación artificial (con las lámparas del laboratorio encendidas). Los videos se encuentran en formato RGB-D de 640X480 pixeles.

2.7 Clasificación

De acuerdo a lo revisado en el estado del arte de reconocimiento de expresiones faciales (FER) se decidió realizar la clasificación con los siguientes algoritmos, ya que fueron con los que se obtiene una mayor precisión al momento de clasificar expresiones faciales:

- Naïve Bayes.
- Máquinas de soporte vectorial (MSV).
- J48.

2.7.1 Naïve Bayes [24]

Naïve Bayes es uno de los algoritmos de aprendizaje más eficientes y efectivos para el aprendizaje automático y minería de datos, este algoritmo probabilístico es uno de los más utilizados por su simplicidad y su capacidad de aprender rápidamente,

que calcula un conjunto de probabilidades contando la frecuencia y combinaciones de valores de un conjunto dado. El algoritmo usa el teorema de Bayes el cual fue creado por Thomas Bayes y se utiliza para revisar probabilidades previamente calculadas cuando se posee nueva información. Es una técnica de clasificación supervisada que tiene como objetivo principal el construir un modelo de clasificación a partir de una base de datos de entrenamiento con etiquetas de clase.

Para resolver la problemática, se utiliza la Ecuación (2.1), la cual define la probabilidad de que se dé un suceso habiendo sucedido otro que influye al anterior:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (2.1)$$

Donde:

- A y B son dos sucesos aleatorios.
- P(A) es la probabilidad de que suceda A.
- P(B) es la probabilidad de que suceda B.
- P(B|A) es la probabilidad de que suceda B dado A es verdadero.
- P(A|B) es la probabilidad a posteriori o condicional, probabilidad de que suceda A dado que B es verdadero.

Asumiendo que las probabilidades de que los datos B son independientes de otros valores para la clase A.

Este algoritmo genera un árbol de decisión, que es el modelo más simple de clasificación ya que asume independencia entre todos los atributos dada una clase. Por lo tanto corresponde a un modelo de atributos independientes. En este caso, la estructura de red es fija y solo es necesario aprender los parámetros.

2.7.2 Máquinas de soporte vectorial

Las MSV máquinas de soporte vectorial (SVM, del inglés *Support Vector Machines*) fueron pensadas en los años 90s para resolver problemas de clasificación binaria, actualmente se utilizan para resolver otro tipo de problemas como la clasificación de imágenes, categorización de textos, agrupamiento, entre otras [25].

Las MSV pertenecen a la categoría de clasificadores lineales, puesto que inducen separadores lineales o hiperplanos, ya sea en el espacio original de los ejemplos de entrada, si éstos son separables o cuasi-separables (con ruido), o en un espacio transformado (espacio de características), si los ejemplos no son separables linealmente en el espacio original.

Este método de aprendizaje se centra en la minimización del denominado riesgo estructural. En la que se selecciona hiperplano que separe los ejemplos más cercanos de cada clase (Figura 2.12).

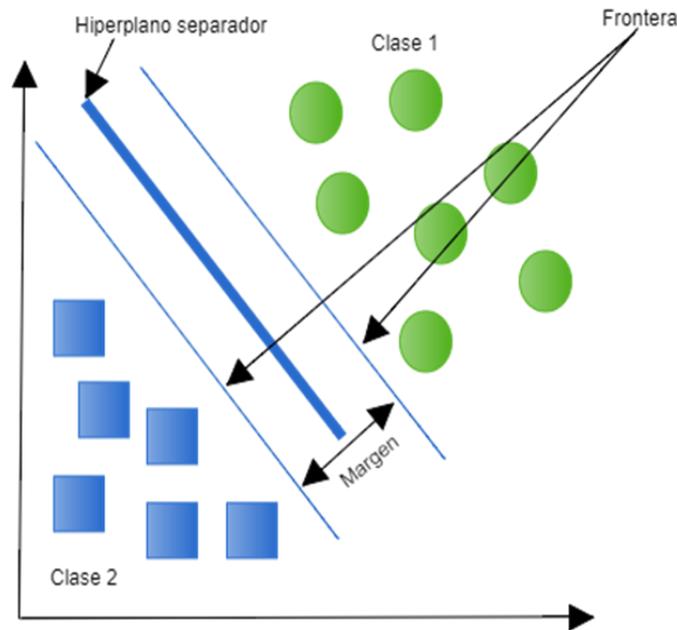


Figura 2.12 Elementos MSV [25].

Una definición kernel asigna a cada par de elementos del espacio de entrada un valor real X correspondiente al producto escalar de las imágenes de dichos elementos en un nuevo espacio F (espacio de características), es decir (Ecuación 2.2):

$$K(x, x^1) \leq \Phi(x), \Phi(x^1) \geq (\phi_1(x)\phi_1(x^1) + \dots + \phi_m(x)\phi_m(x^1)) \quad (2.2)$$

Donde $\Phi: X \rightarrow F$.

Para este proyecto de tesis se utilizó el kernel *RBF* (por sus siglas en inglés, *Radial Basis Function*) o Gaussiano, el cual se encuentra entre los kernel de mapas no lineales, teniendo un espacio dimensional más alto por lo que a diferencia del kernel lineal por lo que puede utilizarse cuando las etiquetas de las clases y los atributos son no lineales [26].

Este kernel requiere de la determinación de dos parámetros C y γ . El primero es una constante que penaliza los errores de clasificación, es decir, los elementos situados en el lado equivocado del hiperplano, y el segundo es un parámetro inversamente proporcional a la amplitud del kernel.

La definición del kernel RBF (Figura 2.3) se puede observar en la siguiente ecuación:

$$K(a, b) = \exp\left(-\frac{(a-b)^2}{2\sigma^2}\right) \quad (2.3)$$

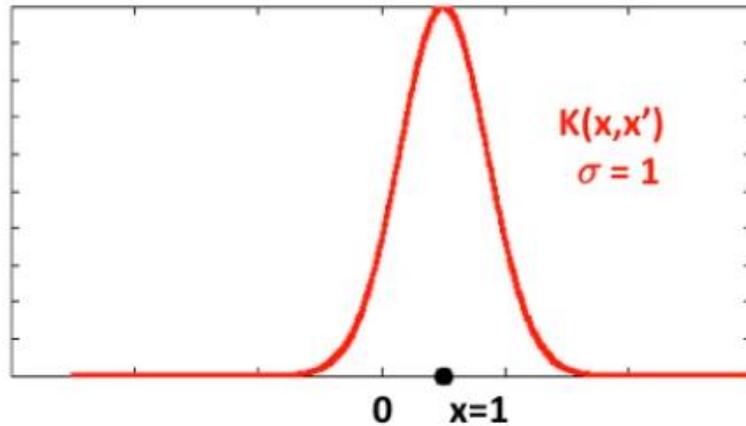


Figura 2.13 Función kernel RBF.

2.7.3 J48

Los árboles de decisión son utilizados en múltiples áreas, tales como la minería de datos, minería de texto, aprendizaje automático, reconocimiento de patrones, etc.

C4.5 es un algoritmo de clasificación el cual en las herramientas *RStudio* y *Weka* es llamado J48. Es un árbol de decisión siendo una forma gráfica y analítica de representar todos los eventos que pueden surgir a partir de una asumida en cierto momento, ayudando a tomar la decisión más acertada desde un punto de vista probabilístico [27].

Este algoritmo es una extensión de ID3, utiliza un enfoque de *divide and conquers* (divide y conquista):

a) Construcción de un árbol J48:

- Seleccionar el atributo para el nodo raíz y crear una ramificación para cada atributo posible.
- Dividir instancias en subconjuntos.
- Repetir recursivamente para cada rama.

b) Entropía

La entropía es una medida del desorden de datos, la cual también es llamada medición de la incertidumbre en cualquier variable aleatoria. La entropía para cualquier P se puede calcular como en la Ecuación 2.4:

$$Entropía(p) = -\sum_{j=1}^n \frac{|p_j|}{|p|} \log \frac{|p_j|}{|p|} \quad (2.4)$$

La entropía condicional es (Ecuación 2.5):

$$Entropía(j|p) = \frac{|pj|}{|p|} \log \frac{|pj|}{|p|} \quad (2.5)$$

La ganancia de información se usa para medir la asociación de entradas y salidas. Es un cambio de estado a estado en entropía de la información. Finalmente, la ganancia de información puede calcularse como en la Ecuación 2.6:

$$Ganancia(p, j) = Entropía(p) - Entropía(j|p) \quad (2.6)$$

Con esta técnica se construye un árbol de decisión como modelo de clasificación, a partir de una base de datos de entrenamiento, una vez construido puede ser aplicado en la base de datos de la cual proporcionara los resultados de la clasificación.

El resultado de este algoritmo puede representarse como reglas de producción si<condiciones>→entonces<conclusiones, acciones, hipótesis>.

2.8 Herramientas para clasificación

Para este trabajo se utilizaron dos herramientas para clasificación:

2.8.1 Weka

Weka es una colección de algoritmos de aprendizaje automático para tareas de minería de datos. Los algoritmos se pueden aplicar directamente en un conjunto de datos o llamar desde su propio código Java. *Weka* contiene herramientas para el procesamiento de datos, clasificación, clustering, reglas de asociación y visualización [3].

También es adecuado para desarrollar nuevos esquemas de aprendizaje automático.

Weka es un software de código abierto emitido bajo la Licencia Pública General GNU (por sus siglas en inglés, *General Public License*).

La interfaz de explorador (*Explorer*) dispone de varios paneles que dan acceso a los componentes principales:

- **Preprocess:** se pueden importar datos desde una base de datos, un CSV, etc., y preprocesar estos datos utilizando algoritmos de filtrado los cuales pueden eliminar registros o atributos según ciertos criterios previamente especificados.
- **Classify:** permite aplicar diferentes algoritmos de clasificación, así como calcular la exactitud del modelo predictivo resultante, además de tener utilidades como visualizar modelos propios como un árbol de decisión.

- **Associate:** muestra las reglas de asociación aprendidas que intentan identificar todas las relaciones entre los atributos y los datos.
- **Cluster:** se tienen acceso a las técnicas de agrupamiento.
- **Selected attributes:** proporciona algoritmos para identificar los atributos más predictivos en un conjunto de datos.
- **Visualize:** muestra un *scatterplot* (matriz de puntos dispersos) donde se puede analizar cada uno de los puntos.

2.8.2 RStudio

R es un lenguaje de programación interpretado, de distribución libre bajo licencia de GNU, y se mantiene en un ambiente de cómputo estadístico y gráfico. En **R** se pueden aplicar técnicas estadísticas [28].

La funcionalidad de **R** consta de paquetes modulares, el paquete básico que se requiere para su ejecución y la mayoría de las funciones fundamentales. Los otros paquetes contenidos en la “base” del sistema incluyen a *utils*, *stats*, *datasets*, *graphics*, *grDevices*, *grid*, *tools*, *parallel*, *compiler*, *splines*, *tcltk*, *stats4*.

La capacidad de **R** es muy sofisticada y mejor que la mayoría de los paquetes estadísticos. **R** cuenta con varios paquetes gráficos especializados. Cabe señalar que, además del paquete básico de **R**, existen 4000 paquetes más.

RStudio es un entorno de desarrollo integrado para el lenguaje de programación **R**, es muy útil para el trabajo interactivo, pero también es un poderoso lenguaje de programación para desarrollo de nuevas herramientas y al ser de software libre lo hace un lenguaje más atractivo [4].

2.9 Discusión

Dentro del estado del arte referente a reconocimiento de expresiones faciales se ha utilizado en los últimos años el sensor Kinect con la librería *facetracking*, la cual contiene la máscara *Candide-3* para el seguimiento del rostro. También descubrió el funcionamiento del sistema de reconocimiento mediante dos diferentes

esquemas, uno para el entrenamiento con una base de datos creada para el mismo con 30 personas y otro para las pruebas el cual fue evaluado con otra base de datos la cual contiene sesiones de video de 15 personas diferentes a las anteriores. También se describieron los 3 algoritmos de clasificación utilizados para el análisis de la base de datos.

Capítulo 3: Juegos serios

La motivación juega un papel muy importante en cualquier rehabilitación. En muchas ocasiones, los pacientes están desmotivados ya que su recuperación consta de varias sesiones, en las cuales las actividades son monótonas e intensivas, que se hacen tediosas después de cientos de repeticiones.

Alentar y motivar a los pacientes para poner esfuerzo en la terapia física es de suma importancia. Una forma de motivación es el uso de juegos para la rehabilitación, lo cual requiere en su diseño la aportación de múltiples áreas como la fisioterapia, bioingeniería y juegos de campo.

Los juegos serios son herramientas interactivas las cuales tienen como objetivo principal desarrollar una habilidad ya sea nueva o perdida, y no sólo la diversión y entretenimiento aunque las engloban, de tal manera que son una solución cuando se busca motivar y alentar al paciente a culminar su recuperación.

En este capítulo se muestra el estado del arte relacionado con juegos serios para la rehabilitación, además del diseño y desarrollo del juego serio propuesto para el sistema integral de rehabilitación.

3.1 Estado del arte referente a juegos serios

Los autores de este artículo son los primeros en construir una interfaz gráfica para un dispositivo de rehabilitación del tobillo, el cual puede ser monitoreado por el

terapeuta. El paciente puede realizar sus ejercicios interactuando con un ambiente virtual, el cual puede ser utilizado desde su hogar.

Este proyecto minimiza las deficiencias, tales como:

- **Fuerza:** ejercicios simulando fuerza resistiva (peso opuesto al movimiento del tobillo).
- **Flexibilidad:** puede mejorar la fuerza mediante movimientos repetitivos acercándose a los límites cercanos de movimiento del tobillo.
- **Balance y coordinación:** son reducidos estos déficits mediante la realización de ejercicios, mejorando el sentido de orientación.

Los pacientes interactúan en el ambiente virtual con un rehabilitador de 3 grados de libertad, donde se les da la instrucción de realizar el movimiento de dorsiflexión-plantarflexión o inversión-eversión, mientras que en pantalla se puede observar una pierna y un pie representando los movimientos que se deben de hacer. La herramienta para el desarrollo de este trabajo fue *WorldToolkit*.

Como resultados se tiene que los pacientes aprendieron a usar fácilmente la interfaz, lo cual los motivó a realizar su rehabilitación; además, tienen la oportunidad de practicar los movimientos antes de hacer la colección de datos que se enviarán al fisioterapeuta.

Serious Games for the Pediatric Anklebot [29]

El objetivo primordial de los autores es brindar un ambiente interactivo para niños entre 5 y 8 con parálisis cerebral con rehabilitación asistida, por medio de un robot y juegos que son controlados mediante los movimientos de dorsiflexión-plantarflexión y eversión-inversión, obteniendo a cambio mayor coordinación, percepción y velocidad del paciente.

Se desarrollaron tres juegos:

1. ***Shipwreck***: en este juego se debe evitar que el barco choque con las rocas que aparecen en pantalla, esto se lleva a cabo mediante los movimientos de dorsiflexión-plantarflexión y eversión-inversión.
2. ***The race of Noah's Ark***: en este juego, el niño tiene que participar en una carrera, en la cual mediante el movimiento de dorsiflexión-plantarflexión correrá, y con los movimientos de eversión e inversión recogerá los animales que encuentre en el camino y evitando los charcos de agua.
3. ***Soccer 2014***: el paciente debe correr hacia el balón (con el movimiento de dorsiflexión-plantarflexión) y defender su portería (movimiento eversión-inversión) de la computadora.

Los autores y desarrolladores de este trabajo tienen como conclusión que es adecuado implementar este tipo de herramientas para los niños que sufren de parálisis cerebral, teniendo un potencial atractivo e incluso adictivo para la rehabilitación.

Serious games for physical rehabilitation: designing highly configurable and adaptable games [30]

El objetivo de este trabajo es brindar una rehabilitación de parálisis cerebral a través de juegos serios en tercera dimensión.

Se desarrollaron cuatro juegos construidos con las herramientas de desarrollo *XML*, *Kinect SDK*, *XNA*, *C++* y *C#*:

1. Simulador de vuelo: es un juego de evadir obstáculos el cual requiere que el jugador controle una nave en dirección horizontal para no chocar con 2 tipos de objetos (rocas y estrellas).
2. Tirar cajas: se debe de lanzar una pelota mediante movimientos horizontales, para tirar las cajas que se encuentran de frente.
3. *Wipeout*: se usa con movimientos horizontales y, al ejecutarlos, se limpia la pantalla descubriendo una imagen en el fondo.
4. *Pickthemup*: el paciente controla dos guantes, con uno de ellos deberá cargar una canasta y con el otro recoger la mayor cantidad de champiñones posible.

A diferencia de otros juegos este puede trabajar con diferentes articulaciones, el terapeuta debe de elegir cual es la que se desea rehabilitar. Al usar estos mini juegos el paciente se siente motivado realizando su terapia.

Robotics and gaming to improve ankle strength motor control and function in children with cerebral palsy- A case study series [31]

El objetivo principal de este trabajo fue investigar la factibilidad de un juego basado en un robot de tobillo para la rehabilitación de niños que sufren de parálisis cerebral. Esta rehabilitación ayuda a los mismos a tener una mejor calidad de vida, demostrando que el uso de las tecnologías brinda más beneficios que la rehabilitación tradicional.

Los juegos que desarrollaron en este trabajo se basan en GMFM por sus siglas en inglés *Gross Motor Function Measure* (que es la medida de la función motora gruesa), se evalúa en cinco dimensiones:

- Acostarse y rodar.
- Sentarse.
- Arrastrarse y de rodillas.

- Pararse.
- Caminar.

Desarrollaron dos juegos: el primero es acerca de un avión que tiene que pasar por ciertos obstáculos mediante el movimiento de dorsiflexión-plantarflexión, en el segundo se debe de ir rompiendo una pared a través de rebotar una pelota con una barra de movimiento (la cual tiene movilidad con el movimiento de eversión-inversión).

Los pacientes, que se sometieron a prueba, realizaron 36 sesiones (3 por semana); al final se obtuvo una mejora en la fuerza, movilidad y rapidez aumentando la calidad de vida de los niños.

Serious Games for Training Rehabilitation and Workforce Development [32]

En este artículo se muestra una recolección de trabajos de la Universidad Clemenson para entrenamiento y rehabilitación. Los juegos multidisciplinarios más significativos desarrollados en esta universidad son los siguientes:

1. ***Patient safety training in a medical simulation:*** este juego es para estudiantes de medicina, mediante esta simulación ellos deben cuidar de sus pacientes tomando las decisiones pertinentes para que ellos mejoren.
2. ***Duck duck punch:*** es para rehabilitación de accidentes cerebrovasculares (*stroke*). El juego está desarrollado en 3D y es basado en el clásico juego de feria de tirar a los patos, es utilizado a través de Kinect, el cual detecta los movimientos del paciente haciendo que los patos caigan.
3. ***Rapid response training system:*** ayuda a practicar a los enfermeros interactuando con los pacientes. El enfermero debe determinar que paciente es el que se está deteriorando más para ayudarlo a sobrevivir.
4. ***Virtual e-school for aviation and automotive technology education:*** es para estudiantes de aviación, ayudando a interactuar como si estuvieran en una simulación real, manipulando los calibradores y micrómetros.

Por trabajos como estos se puede observar que los juegos serios no tienen como única área la rehabilitación, sino también el capacitar a algunos estudiantes en su rama, entre muchas otras aplicaciones.

Serious games for stroke rehabilitation employing immersive user interfaces in a 3D virtual environment [33]

Los autores de este trabajo impartieron protocolos de rehabilitación que pudieran ser utilizados desde el hogar del paciente que ha sufrido de un accidente

cerebrovascular, mientras el fisioterapeuta puede observarlo desde su consultorio. Esto con ayuda del registro personal de salud PHR (por sus siglas en inglés *Personal Health Register*).

Se realizaron varios juegos y un avatar del paciente para que este pueda ver los ejercicios que se están haciendo. Los fueron desarrollados con *MS Kinect SDK*, *Open NI drivers*, *Unity3D*, *Brekel/Auto Motion Builder*. Los mini juegos son:

- Memoria.
- Destruir letras.
- Mover objetos.
- Realizar los movimientos en pantalla.

Después de realizar las pruebas, demostraron que el poder realizar su rehabilitación en casa es muy efectivo ya que los pacientes se desenvuelven más en un ambiente familiar. Además, económicamente se gasta menos ya que se requiere en menor proporción la presencia del terapeuta que en una rehabilitación tradicional.

Serious games for assessment and rehabilitation of ankle movements [34]

Los autores tuvieron como objetivo el desarrollar juegos serios para la rehabilitación del tobillo usando un dispositivo robótico, ya que esto puede servir como una terapia virtual, evaluando los rangos de movimiento de dorsiflexión-plantarflexión y de su fuerza muscular.

Se desarrollan dos juegos en *C#* y *XNA* los cuales fueron integrados con un robot para el tobillo:

1. **“O Guloso”**: el cual fue diseñado para evaluar el rango de movimiento del tobillo. En este juego el paciente debe de tomar la comida que sale en pantalla mediante el movimiento de dorsiflexión-plantarflexión.
2. **“O Atleta”**: se debe de saltar los obstáculos en pantalla, mediante este juego se puede evaluar el esfuerzo de torsión.

Se hicieron pruebas a 19 personas sanas y 19 con hemiparesia (parálisis de un lado del cuerpo), algunos de ellos con lesiones del lado derecho y otros del lado izquierdo (a los cuales les costó más trabajo el llevar a cabo los juegos).

El resultado mostró que puede ser una muy buena alternativa para la rehabilitación del tobillo, la cual puede brindar rehabilitación asistida y pasiva, ya que es muy sencillo de utilizar.

The mobile RehApp™: an AR-based mobile game for ankle sprain rehabilitation [35]

RehApp™ es una aplicación interactiva para la rehabilitación del tobillo, la cual está desarrollada en realidad aumentada. En ella se realizan dos juegos:

Mediante los movimientos del tobillo el paciente debe de formar las letras del abecedario.

El segundo juego se llama “el limpiaparabrisas” el cual consiste en mover el pie de un lado al otro hasta descubrir la imagen oculta.

Como conclusión, los autores mencionan que al usar este tipo de dispositivos hace que la rehabilitación sea un proceso más fácil, practico y portable (ya que se puede realizar desde su hogar) por lo que se vuelve un proceso autónomo, motivando a los pacientes a realizar sus ejercicios para la recuperación del tobillo.

A virtual-reality tracking game for use in robot assisted ankle rehabilitation [36]

Zhang en este trabajo realizo un juego de realidad virtual asistido por un robot, mediante el cual sea más fácil llevar a cabo los ejercicios de rehabilitación de tobillo de una manera entretenida, interesante e interactiva eliminando la aburrición.

El juego desarrollado trata de hacer coincidir al punto negro con el azul. Para ello el paciente pone su pie en el rehabilitador y utiliza los movimientos del tobillo para hacer llegar el punto azul al mismo lugar que el negro, esto se realiza durante 10 rounds y al finalizar se muestra la puntuación final del paciente.

Al realizar las pruebas, los pacientes comentaron si había más juegos divertidos como éste, y que deseaban obtener una mejor puntuación la siguiente vez que lo jugaran. Llegando a la conclusión de que un juego para rehabilitación es atractivo y una buena alternativa para evitar el aburrimiento al realizar ejercicios repetitivos.

Designing motivational games for stroke rehabilitation [37]

El objetivo de este trabajo fue desarrollar juegos para la rehabilitación de un accidente cerebrovascular (*stroke*) por medio de la motivación. En total se desarrollaron tres juegos basándose en SCRIPT por sus siglas en inglés *Supervised Care and Rehabilitation Involving Personal Telerobotics*:

1. **Marble maze:** en este juego el paciente debe de hacer embonar una pelota en el hueco en pantalla, para ello debe deslizarse por un laberinto.
2. **Run Jack run:** es una carrera en la que el paciente debe de evitar los obstáculos que tenga frente a él.

3. **Burger hub:** el paciente juega a preparar hamburguesas, desde cortar las verduras hasta colocar todos los ingredientes en su posición.

Las herramientas utilizadas en este trabajo para el desarrollo de los juegos fueron *Unity3D* y *C#*. Como conclusión se encontró que el uso de estos juegos para rehabilitación motiva a los pacientes y los motiva a jugar largas jornadas.

Development of 2-DOF Ankle Rehabilitation [38]

Pretende brindar un sistema que ayude a mejorar el tobillo de lesiones ortopédicas y neurológicas. *The Virtually-Interfaced Robotic Ankle and Balance Trainer* (Vi-Rabt, entrenador de balance y dispositivo robótico para el tobillo virtualmente interconectado) es un sistema de rehabilitación diseñado para proporcionar terapia pasiva, activa, asistida o resistiva para el tobillo.

El juego desarrollado por los autores es una especie de laberinto circular en el que el paciente, por medio de los movimientos del tobillo, debe de llegar hasta el centro. Como herramienta de desarrollo fue utilizado *Labview*.

Este juego fue probado con un paciente sano obteniendo resultados muy alentadores, lo que indica que puede ser probado en ensayos clínicos.

Development of a Complete Game Based System for Physical Therapy [39]

Los autores de este trabajo buscaban el brindar rehabilitación física mediante una herramienta interactiva como lo puede ser un juego. Para el desarrollo de estos juegos se cree necesario que deben de contar con las siguientes características:

- Adaptabilidad.
- Supervisión.
- Portabilidad.
- Interfaz.
- Género (el cual se refiere a RSM por sus siglas en inglés Repetition of Specific Movements (repetición de movimientos específicos) y RRO por sus siglas en inglés Reach Random Objects (alcanzar objetos aleatorios)).

Se desarrollaron dos juegos utilizando las herramientas *XNA game engine*, *Kinect Windows SDK*, *Unity3D*, *Kinect Wrapper Package* (para *Unity3D*), los juegos son los siguientes:

1. **Asteroids kRehab:** en el que el paciente debe realizar repeticiones de movimientos específicos definidos por el fisioterapeuta, la dinámica es controlar una nave espacial evitando los asteroides, esto se realiza mediante movimientos de brazos.

2. **3D pong kRehab:** este juego puede ser para dos personas y trata de interactuar con un juego de ping pong virtual, en el que las manos serán las raquetas de los jugadores.

Como conclusión se tiene que es importante el brindar motivación al paciente para realizar su rehabilitación para mejorar su condición, por lo que el uso de los juegos serios se considera una buena herramienta.

Vision-based interaction as an input of serious game for motor rehabilitation [40]

En este trabajo se tuvo como objetivo el demostrar los beneficios de un sistema basado en visión artificial para adultos con parálisis cerebral.

El sistema fue diseñado para visualizar la postura y balance del paciente, el cual está frente a la pantalla en simulación espejo, lo que lo ayuda a interactuar con el juego usando sus propios movimientos.

Los autores creen que un sistema integral como el que proponen debe de contar con las cuatro características siguientes:

- Retroalimentación.
- Adaptabilidad.
- Elementos motivacionales.
- Supervisión.

Las herramientas de desarrollo que utilizaron son *C++*, *OpenNI* por *Kinect*, *OpenCV* y *Q+*.

Serious games for functional rehabilitation [41]

Los autores se plantearon el desarrollar un juego serio en tiempo real, ya que se tiene la idea de que una de las soluciones para tener una rehabilitación más eficiente es hacerlo más divertido y atractivo para tener una mejor interacción paciente-sistema de rehabilitación, de tal manera también pueda ser utilizado, así como en la clínica, en el hogar.

El juego tiene ejercicios locomotores para la rehabilitación músculo-esquelética. Estos ejercicios son 6 diferentes los cuales están divididos en niveles con distintas dificultades. Dichos movimientos son:

- Movimiento de cadera.
- Saltar.
- Levantamiento de pierna.
- Postura de una pierna.

- Movimiento de rodilla.
- Sentadillas.

Las herramientas utilizadas para la implementación del juego fueron *Microsoft XNA Game Studio*, *Visual Studio .Net*, *C#*.

Se puede observar que el utilizar juegos es una buena solución para la rehabilitación musculo esquelética; Además de que se construyó un avatar en 3D que simulaba lo que el paciente hacía, ayudando a retroalimentarse mediante la vista espejo.

3-DOF admittance control robotic arm with a training of hemiparetic hand [42]

Los autores desarrollaron un juego virtual 3D, el cual puede ser ejecutado a través de un brazo robótico de 3 grados de libertad, el cual sirve como rehabilitación de parálisis de mano.

El juego trata de reventar las pelotas que aparecen en pantalla. Cuando el paciente llega a 5 puntos (1 por cada pelota) gana, obteniendo una retroalimentación en la que se indica el tiempo en que fue realizado.

Para el control y desarrollo del juego fueron utilizadas las herramientas de *Matlab*, *RoboPlus toolbox*, *Unity3D*.

Serious games development for ankle rehabilitation aiming at user experience [43]

Los autores de este artículo pretenden hacer una mejora a un juego creado por IMT por sus siglas en inglés *Interactive Motion Technologies* perfeccionando la estética y el juego. Dicho juego es controlado mediante el “anklebot” creado por IMT, también tenían como objetivo el motivar al paciente para tener una mejor experiencia. Para la rehabilitación del paciente se llevó a cabo el siguiente procedimiento:

1. Juego libre: no se tienen restricciones ni ayuda para el movimiento, mediante esto se puede evaluar el rango de movimiento del tobillo que tiene el paciente en ese momento.
2. Asistido: con ayuda para el movimiento, pero sin restricciones, lo que ayuda a evaluar el comportamiento de la fuerza del tobillo.
3. Resistivo: con restricción y sin ayuda para el movimiento, lo que ayuda a probar la fuerza actuando como una restricción.

Se realiza un juego inspirado en el naufragio de IMT, el cual es como un ping pong doble, y puede moverse en el eje x mediante el movimiento de eversión-inversión y

en el eje y con el de dorsiflexión-plantarflexión. Este juego fue elaborado con la herramienta de desarrollo *Unity3D*.

Su evaluación se realizó primero con personas sanas, lo que resultó en éxito; posteriormente se probó con pacientes lesionados. También es importante mencionar que cuando el paciente alcanza cierto desplazamiento establecido, el juego, se ajusta automáticamente teniendo un nuevo límite de movimiento, ayudado a mejorar el rango de movimiento del paciente.

3.2 Análisis del estado del arte referente a juegos serios

Los juegos serios necesitan ser interesantes, entretenidos e interactivos [36], el tener un concepto interesante apoyará la motivación del paciente. También es recomendable considerar tener una interfaz visual simple y un control sencillo para que el periodo de aprendizaje sea corto [29], ayudando al paciente a ser autónomo y de ser posible que pueda realizar la terapia desde su hogar.

Cuando se trata de rehabilitación de tobillo, la terapia que se debe realizar es por medio de la repetición de movimientos específicos (RSM) como lo pueden ser los movimientos de dorsiflexión-plantarflexión o eversión-inversión [44] [29] [31] [34] [35] [36] [38] [39] [43].

Los artículos revisados permiten ver que en los últimos trabajos han utilizado la retroalimentación espejo, en la que mediante un avatar el paciente puede observar los movimientos que realiza en la pantalla, facilitando la interacción con el juego [40].

La Tabla 3.1 muestra un resumen de los trabajos revisados en el estado del arte, los juegos serios pueden ser utilizados en múltiples áreas como lo puede ser la capacitación en algún ámbito laboral o el aprendizaje, este trabajo estará enfocado al área de rehabilitación y específicamente del tobillo.

El juego desarrollado en esta tesis es dirigido a personas jóvenes, ya que se cree que son los que más atención y motivación necesitan para no abandonar su proceso de rehabilitación; además que los juegos harán su terapia más atractiva y divertida.

Tomando en cuenta las herramientas de desarrollo reportadas, se eligió *Unity3D* como herramienta para la implementación del juego, ya que es una plataforma de desarrollo flexible que puede ser implementada en todos los sistemas operativos (Linux, Mac, Windows, etc.).

El juego que se desarrolló contempla sólo el movimiento de dorsiflexión-plantarflexión, ya que el rehabilitador de tobillo con el que se integra es de un grado de libertad.

Tabla 3.1 Resumen del estado del arte referente a juegos serios.

Autor	Área	Rehabilitación	Dirigido a:	Portabilidad	Interfaz	Herramientas de desarrollo	No. de juegos	RSM	Retroalimentación
Girone (1999) [44]	Terapia física	Esguince de tobillo	N/A	Hogar	Realidad virtual	WorldToolkit	1	D-P E-I	Sí
Michmizos (2012) [29]	Terapia física	Tobillo por parálisis cerebral	Niños	Clínica	2D	N/A	3	D-P E-I	Sí
Omelina (2012) [30]	Terapia física	Parálisis cerebral	Niños	Hogar	3D	XML, C#, Kinect SDK, XNA, C++	4	No	Sí
Bertrand (2013) [32]	Múltiples áreas	Solo en un juego para accidente cerebrovascular	Estudiantes o pacientes	N/A	3D	N/A	4	No	Sí
Burdea (2013) [31]	Terapia física	Tobillo por parálisis cerebral	Niños	Clínica y hogar	Realidad virtual	N/A	2	D-P E-I	Sí
Vogiatzaki (2013) [33]	Terapia física	Accidente cerebrovascular	N/A	Hogar	3D	OpenNI, Unity3D, Brekel	4	No	Sí
Gonçalves (2014) [34]	Terapia física	Tobillo por hemiparesia	N/A	N/A	2D	C#, XNA	2	D-P	Sí
García (2014) [35]	Terapia física	Esguince de tobillo	N/A	Hogar	Realidad virtual	N/A	2	D-P E-I	Sí
Zhang (2014) [36]	Terapia física	Esguince o fractura de tobillo	N/A	Clínica	Realidad virtual	N/A	1	D-P	Sí
Shah (2014) [37]	Terapia física	Accidente cerebrovascular	N/A	Hogar	3D	Uniry3D, C#	3	No	Sí
Farjadian (2014) [38]	Terapia física	Tobillo por accidente cerebrovascular o parálisis cerebral	Adultos	N/A	2D	Labview	1	D-P E-I	Sí
Menezes (2014) [39]	Terapia física	Accidente cerebrovascular	N/A	Clínica y hogar	3D	XNA, Kinect SDK, Unity3D	2	Solo un juego D-P	Sí
Jaume.i-Capó (2015) [40]	Terapia física	Parálisis cerebral	Adultos	N/A	2D	C++, OpenNI, OpenCV, Q+	1	No	Sí, además retroalimentación espejo
Tannous (2015) [41]	Terapia física	Locomotora	Adolescentes	Clínica y hogar	3D	XNA, C#, Visual studio .Net	1	No	Sí, además retroalimentación espejo
Agas (2015) [42]	Terapia física	Parálisis cerebral	N/A	Clínica	3D	Matlab, RoboPlus, Unity3D	1	No	Sí
Pasqual (2016) [43]	Terapia física	Tobillo	N/A	N/A	3D	Uniry3D, C#	1	D-P E-I	Sí

Donde: D-P = Dorsiflexión-Plantarflexión, E-I=eversión inversión, RSM= Repetición de movimientos específicos

3.3 Diseño y desarrollo del juego serio

En el diseño del juego serio desarrollado se consideraron los siguientes puntos:

- Su funcionalidad corresponde a un grado de libertad.
- Es posible ser implementado con el rehabilitador de tobillo, el cual es retroalimentado a partir de un sistema de reconocimiento de expresiones faciales.
- Contiene tres niveles de dificultad.

Al ser un juego serio pretende brindar una habilidad, en este caso una que fue perdida debido a alguna lesión. Las habilidades que se recuperan al final de las sesiones de rehabilitación son la fuerza y el movimiento del tobillo. Consultando un fisioterapeuta se determinó que para llegar a obtener dichas habilidades los niveles deben de estar divididos en 3, primero se debe de tener una fase analgésica en la que los pacientes toman medicamento para eliminar el dolor y la inflamación, posteriormente se puede iniciar recuperando la fuerza de la articulación del tobillo comenzando con movimientos de dorsiflexión-plantarflexión y por último recuperando la movilidad realizando los movimientos de tobillo con un rango de movimiento mayor cada vez (Figura 3.1)²:



Figura 3.1 Estructura de los niveles del juego serio.

El juego consiste en que el personaje principal tiene que saltar los obstáculos que vayan pasando en la pantalla. El paciente, para saltar, debe realizar el movimiento de plantarflexión (flexionar el pie hacia abajo) y al regresar su pie a la posición original realizará el movimiento de plantarflexión.

En el nivel 1, se pretende empezar a recuperar la fuerza perdida del tobillo, ya que previamente el paciente estuvo en una etapa analgésica en la cual tomó medicamentos y no tuvo movimiento ningún tipo de movimiento en la articulación, por lo que debe de empezar a moverlo poco a poco. Por este motivo, en el primer nivel serán puestos muy pocos obstáculos, los cuales estarán separados a 6 segundos uno del otro ya que se tiene un rango de movilidad mínimo.

² Lic. Guillermo Gutiérrez Vásquez, Licenciado en terapia física, Escuela de terapia física DIF del estado de Chiapas.

Durante el nivel 2 se espera una mejoría del paciente, por lo que pretende mejorar la movilidad. El juego en este nivel contiene obstáculos semejantes a los encontrados en el nivel 1, pero se presentaran con mayor frecuencia, iniciando con la separación de 3 segundos cada uno, teniendo como fin que el paciente pueda tener un mayor grado de movilidad.

Por último, en el nivel 3 el paciente está casi recuperado, por lo que el juego presenta dos obstáculos diferentes, uno similar al correspondiente al nivel dos, y otro más amplio, los cuales aparecerán con una frecuencia de 3 segundos. A diferencia del nivel dos es que pueden pasar dos obstáculos al mismo tiempo.

Todos los niveles serán retroalimentados por el sistema de reconocimiento de expresiones faciales, el cual detecta si el paciente se encuentra motivado, desmotivado o con dolor. A partir de esta información, si el paciente se encuentra motivado el nivel permanece en el mismo estado, si se encuentra desmotivado los obstáculos aparecen con mayor frecuencia y si tiene dolor aumenta el tiempo en que pasa un obstáculo con respecto al siguiente.

3.3.1 Diseño del rehabilitador del tobillo de 1 gdl

El rehabilitador utilizado en este proyecto fue desarrollado por Víctor Manuel Penagos Pérez en su tesis titulada Diseño y construcción de un robot paralelo con dos guías lineales [45].

El diseño final del robot paralelo, fue seleccionado después de realizar simulaciones con el programa de Adams View y con base a los resultados obtenidos en su comportamiento cinemático y dinámico se optó por el diseño de la Figura 3.2.

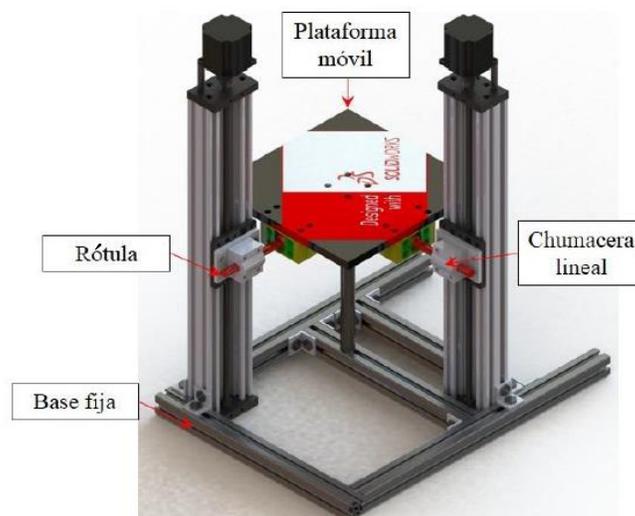


Figura 3.2 Diseño del robot paralelo [45].

El sistema de control se implementa con una tarjeta *arduino Mega* (Figura 3.3). *arduino* es una placa de microcontrolador que se conecta a la computadora por medio de un cable USB, contiene 14 entradas y salidas digitales de las cuales 6 pueden ser usadas como salidas PWM, tienen 6 entradas analógicas en el cual se pueden conectar diversos sensores. En la programación del ciclo de trabajo de cada motor a paso se consideró que una vuelta del motor equivale a 8 mm de carrera del carro de la guía lineal, por lo que al resolver la cinemática del robot paralelo se tiene que para alcanzar una posición de 30 grados se requieren 8 vueltas de cada motor [45].

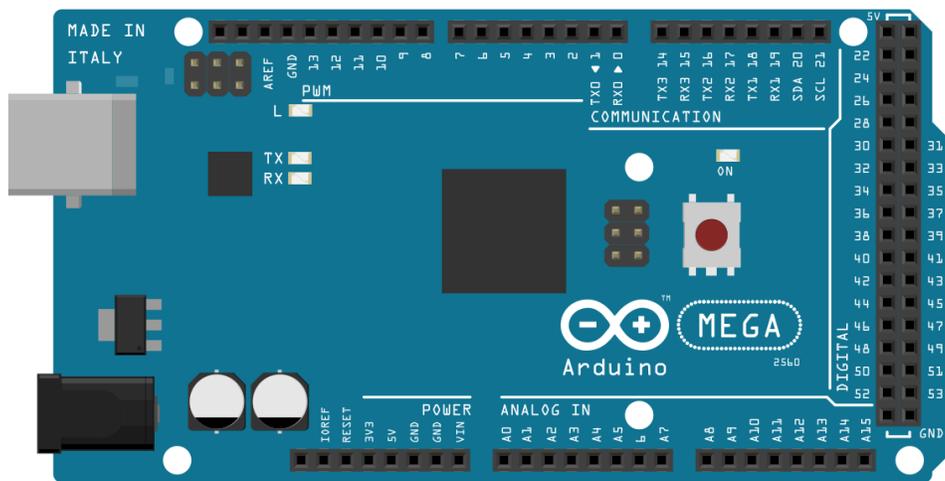


Figura 3.3 Tarjeta arduino Mega.

También se agregó al rehabilitador un sensor el cual detecta la fuerza que ejerce el pie del paciente en rehabilitador al momento de realizar el movimiento de dorsiflexión-plantarflexión (Figura 3.4).

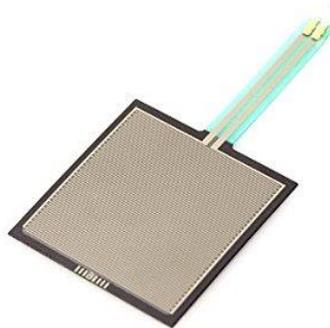


Figura 3.4 Sensor de fuerza.

La tarjeta *arduino Mega* en conjunto del sensor de fuerza es lo que permite que el juego serio sea implementado con el rehabilitador y controlado por el mismo, ya

que a partir de la tarjeta se envían señales que realizan el salto del personaje principal cuando el tobillo hace el movimiento de dorsiflexión-plantarflexión.

En la Figura 3.5 se muestra el prototipo físico:

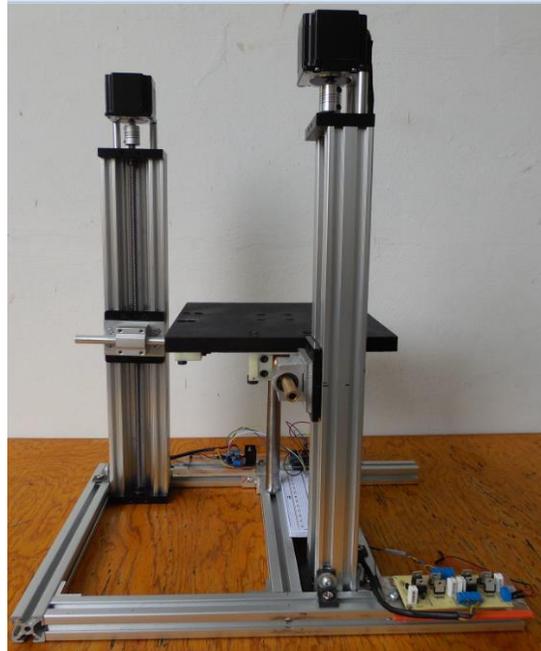


Figura 3.5 Prototipo físico.

3.3.2 Psicología del color en juegos serios

Según Hervás [46] la psicología del color es un aspecto muy importante al realizar juegos ya que a partir de los colores se pueden transmitir emociones a quien esté interactuando (Figura 3.6).

Se decidió utilizar tonalidades azules y verdes ya que transmiten sensación de bienestar, además, se colocaron obstáculos de color rojo teniendo como intención captar la atención de los jugadores creando alerta.



Figura 3.6 Psicología del color.

3.3.3 Animaciones

Al crear las animaciones del personaje se crea un diagrama de estados, para este juego el jugador tiene tres posibles estados los cuales son correr, saltar o morir (Figura 3.7).

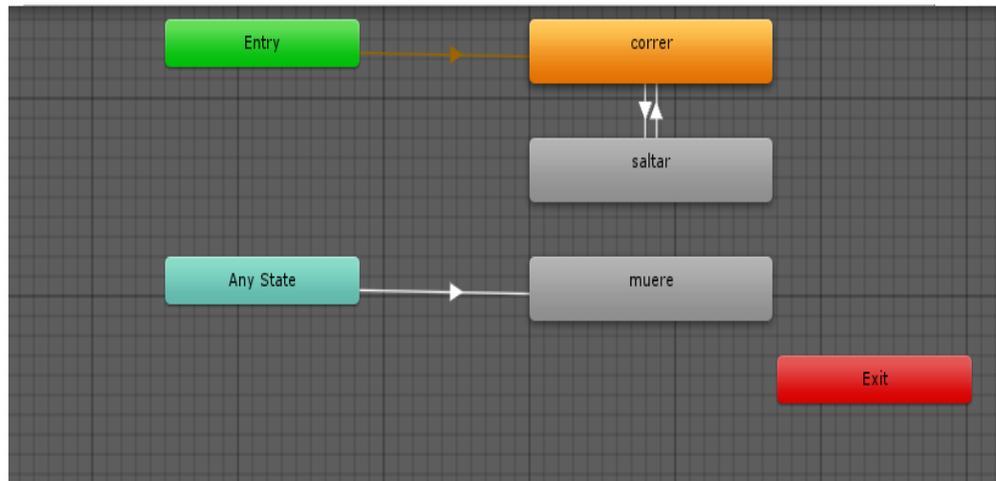


Figura 3.7 Diagrama de estados.

Se puede pasar de un estado a otro mediante las transiciones, para este caso se crearon dos variables, una booleana que indica si el personaje está lastimado (cuando llega a chocar con algún obstáculo), y otra flotante que indica que tiene velocidad. El personaje comienza corriendo y al oprimir una tecla (rehabilitador) salta (Tabla 3.2).

Tabla 3.2 Variables.

De:	A:	Velocidad	Lastimado
Correr	Saltar	Mayor que 0.01	
Saltar	Correr	Menor que 0.01	
Cualquier estado	Morir		Verdadero

Se utilizan tres personajes cada uno para cada nivel del juego, estos fueron elegidos de acuerdo a sus pesos, es decir, el más pesado es utilizado en el nivel 3 ya que necesita de un mayor esfuerzo para poder saltar y viceversa, el más ligero se presenta en el nivel 1 (Figura 3.8).



Figura 3.8 Personajes.

3.4 Discusión

En este capítulo se estudió el estado del arte referente a juegos serios para rehabilitación, por lo que se determinó que se utilizaría la herramienta *Unity* para el desarrollo del juego serio. También se pudo observar el diseño y desarrollo del juego serio para rehabilitación del tobillo el cual se hizo considerando el movimiento de dorsiflexión-plantarflexión correspondiente a 1 gdl. Los tres niveles del juego serio fueron basados en los niveles de rehabilitación del tobillo, en los cuales se recupera primero la fuerza del tobillo para posteriormente recuperar su movilidad.

Capítulo 4: Diseño de la integración

En este capítulo se muestra el diseño de la integración de los tres sistemas (sistema de reconocimiento de expresiones faciales, el juego serio y el rehabilitador de tobillo).

Para el desarrollo e integración de estos tres sistemas fueron utilizadas diferentes herramientas tanto de software como de hardware, las cuales son mencionadas en la Tabla 4.1.

Se utilizó una Laptop con sistema operativo Windows 10, ya que de acuerdo con las especificaciones para utilizar el *sdk* de Kinect es necesario tener un sistema operativo Windows 7 o posterior.

El seguimiento del rostro del paciente se llevó a cabo mediante el sensor Kinect, utilizando la librería *facetracking* en visual *c++*, para el desarrollo del juego serio fue utilizado el motor de videojuegos multiplataforma *Unity engine*.

Mientras que el rehabilitador de tobillo puede realizar de manera activa (el paciente ejerce la fuerza), el movimiento de dorsiflexión-plantarflexión, a través de un sensor de fuerza, todo esto controlado a partir de un arduino Mega, el cual envía la señal al juego serio para que el paciente salte el obstáculo en pantalla.

Tabla 4.1 Herramientas de desarrollo.

HARDWARE	
Laptop con procesador AMD A8-7410 APU con AMD Radeon R5 graphics 2.20 Ghz., 8 gb en RAM y un sistema operativo Windows 10 de 64 bits	
Sensor Kinect de Xbox 360	
Sensor power supply para Kinect	
Rehabilitador de tobillo de 1 gdl (plantarflexión-dorsiflexión)	
Arduino Mega	
Sensor de fuerza	
SOFTWARE	
Kinect sdk: librería <i>facetracking</i> [47]	
Microsoft visual studio Express 2013 (C++) [48]	
MySQL [49]	
Unity engine (C#) [50]	
Arduino [51]	

En la Figura 4.1 se muestra una imagen representativa de la integración del sistema:

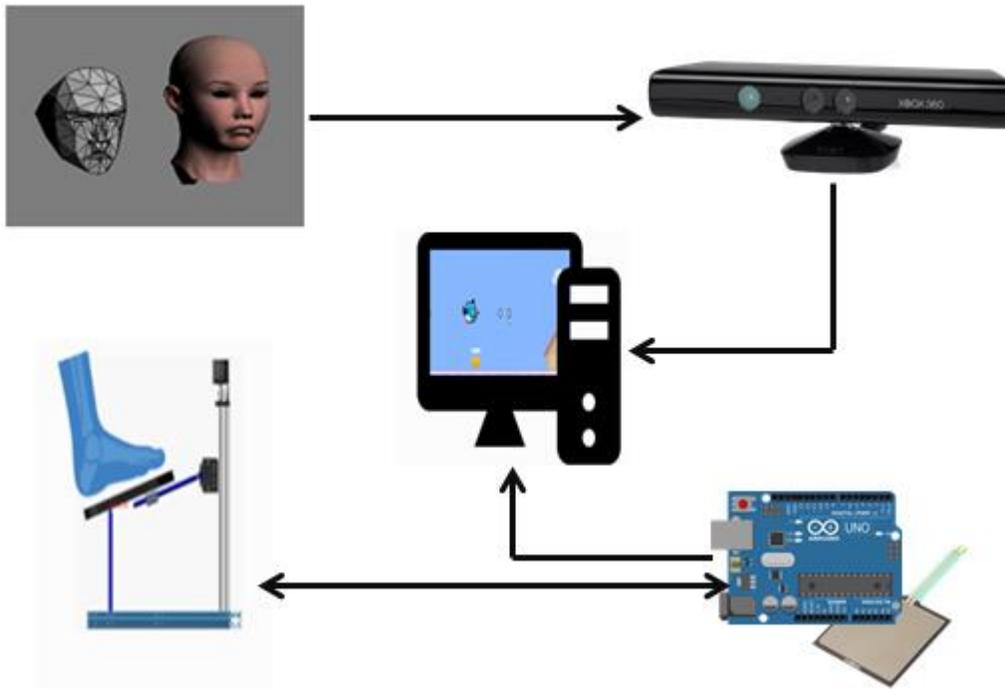


Figura 4.1 Esquema de la integración.

Los sistemas fueron integrados de la siguiente manera:

1. El sistema de visión artificial para el reconocimiento de expresiones faciales está desarrollado en visual *c++*. Este sistema recibe como entrada las imágenes adquiridas por el sensor Kinect; el vector de características de cada imagen es clasificado mediante el modelo J48 (generado fuera de línea mediante el análisis de las bases de datos de entrenamiento y validación), la información resultante (el resultado de la clasificación) es almacenada en una base de datos en *mysql*.
2. El juego serio es retroalimentado mediante el sistema de visión artificial el cual mediante un script de *unity c#* lee cada 10 milisegundos los registros en la base de datos de *mysql* generada en tiempo real. Lo cual según la expresión que tenga el paciente aumenta o disminuye la frecuencia de los obstáculos.
3. A partir del puerto COM de la computadora se implementa la integración del rehabilitador con el juego serio, el cual mediante el sensor de fuerza manda una señal cuando el paciente realiza correctamente el movimiento, haciendo que el personaje principal del juego salte.

En la Tabla 4.2, se muestran los valores considerados para la integración de los sistemas, lo cual hace el salto del paciente en el juego serio. Donde F.O. es la frecuencia de aparición de los obstáculos y F.M.E. es la fuerza mínima ejercida en el rehabilitador para generar el movimiento en el juego.

Tabla 4.2 Valores considerados para la integración.

Nivel del juego serio	F.O. Motivado	F.O. Distraído	F.O. con dolor	F.M.E.	Grado de movimiento
Nivel 1	6 segundos	3 segundos	10 segundos	10 N	-5° a -15°
Nivel 2	3 segundos	2 segundos	6 segundos	15 N	-15.1 ° a -20°
Nivel 3	3 segundos	2 segundos	6 segundos	20 N	-20.1° A -30 °

4.1 Discusión

En este capítulo se mostraron las herramientas de software y hardware para la implementación de los sistemas, así como el diseño de la integración de los tres sistemas (el de reconocimiento de expresiones faciales, juego serio y rehabilitador de tobillo) mediante un arduino Mega que manda la señal de salto al juego serio.

También se pudo observar los valores considerados para la integración los cuales consideran la fuerza mínima ejercida y el grado de movimiento que necesita realizar el paciente sobre el rehabilitador para hacer saltar al personaje principal.

Capítulo 5: Experimentación

En este capítulo se muestran las pruebas realizadas al experimentar con el sensor Kinect, utilizando la librería *facetracking* para el reconocimiento de expresiones faciales del paciente, además de las pruebas para la verificación y validación de la funcionalidad del sistema integral (rehabilitador-juego serio- sistema de visión artificial). También se evalúa el cumplimiento de los objetivos planteados para la realización de este trabajo.

5.1 Ambiente de pruebas

Para la realización de las pruebas fue necesario contar con las siguientes herramientas:

- Laptop con sistema operativo Windows 10.
- Sensor Kinect.
- Rehabilitador de tobillo de 1 gdl.
- Sensor de fuerza.
- *MySQL, Unity Engine C#, Visual C++, Kinect SDK y arduino.*
- *Weka y RStudio.*

5.2 Plan de pruebas

El objetivo de las pruebas es validar el correcto funcionamiento de los sistemas integrados, así como comprobar que se alcanzaron los objetivos planteados, donde

se evalúa el correcto funcionamiento del reconocimiento de las expresiones faciales especificadas y las señales emitidas por el rehabilitador.

Se realizaron 5 pruebas las cuales se mencionan a continuación:

- Prueba 1: Condiciones para el uso de Kinect.
- Prueba 2: Rango de movimiento del tobillo en personas sanas.
- Prueba 3: Comparación de algoritmos de clasificación.
- Prueba 4: Funcionalidad de los sistemas integrados.
- Prueba 5: Integración del sistema en tiempo real.

5.2.1 Prueba 1: Condiciones para el uso de Kinect

Se realizaron tres pruebas para visualizar las condiciones óptimas para el uso de Kinect, para ello se observaron las condiciones de luz, la distancia a la que debe de estar el paciente y si se pueden utilizar accesorios sin que altere el reconocimiento.

a) Iluminación

La luz es un factor importante al realizar capturas con cualquier cámara, por lo que se ha decidido evaluar cuatro condiciones de iluminación diferente:

- Iluminación natural.
- Iluminación artificial.
- Iluminación artificial excesiva.
- Sin luz.

En la Figura 5.1 se muestran las imágenes capturadas de cada una de estas condiciones, evaluadas para observar el desempeño de Kinect, a partir de esto se hizo una comparación de cuál es la mejor condición para utilizar el sistema.

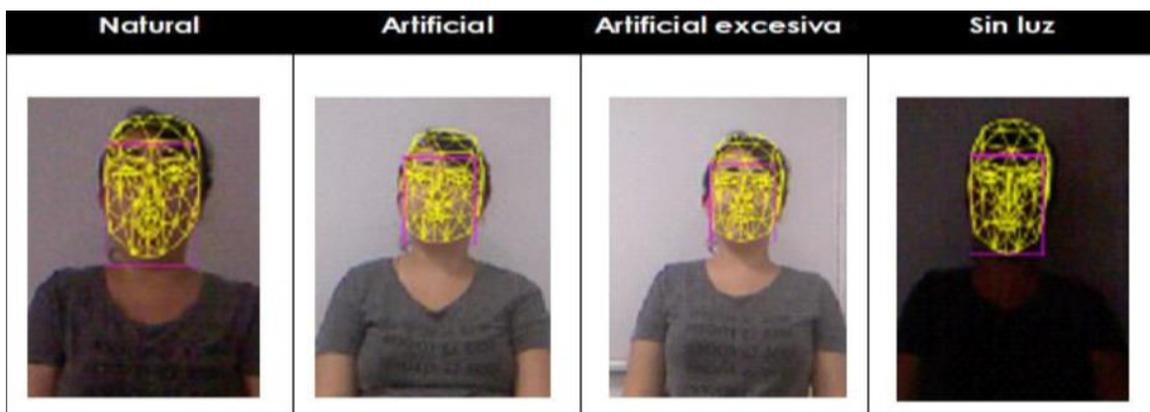


Figura 5.1 Comparación de condiciones de luz.

El sensor Kinect es robusto a las variaciones de iluminación, en la Figura 5.1 se muestran diversos ejemplos en los cuales la respuesta es buena, entonces, a partir

de la captura de estas imágenes se puede ver que a pesar de variación en la iluminación, la máscara candide-3 se posiciona sobre el rostro; sin embargo, lo más apropiado es utilizar una luz artificial que no provoque ningún tipo de sombra en el rostro que podría hacer variar los valores obtenidos.

b) Distancia al sensor

La distancia recomendada para el buen funcionamiento del Kinect es entre los 60 cm a 1.20 m por lo que se realizaron pruebas entre esas distancias, a menores distancias y a mayores. En la Figura 5.2 se muestran los resultados obtenidos.

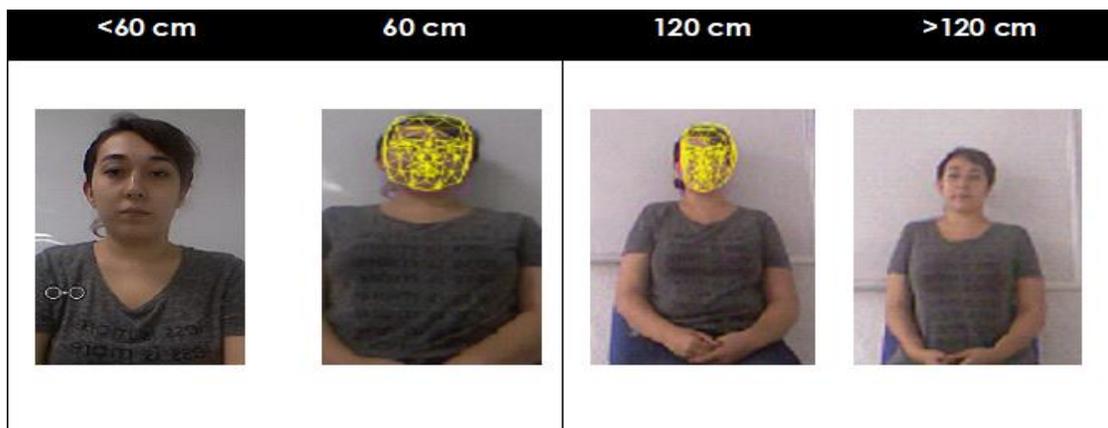


Figura 5.2 Condiciones de distancia.

Como se puede observar se debe tener una distancia entre los 60 y los 120 cm, como lo dice Kinect, ya que a una distancia mayor o menor, no se detecta el rostro y con ello no es posible se posiciona la máscara Candide -3 sobre el rostro para hacer el seguimiento y descripción del mismo.

c) Uso de accesorios

Si una persona presenta algún tipo de accesorio en el rostro, tales como bigote, fleco, lentes, gorras, etc., el sensor Kinect detectará tales objetos y los considerará como parte de los componentes del rostro, induciendo errores en la descripción del mismo. Por ejemplo, en el caso del bigote la máscara candide-3 considera como si fuera el labio superior y provoca que la detección de las comisuras de la boca sean clasificadas como hacia abajo, y con ello el sistema de reconocimiento de expresiones faciales le asignaría la categoría de “con dolor o molestia”. En cuanto al fleco, los lentes y las gorras, la máscara candide-3, considera como si las cejas estuvieran pegadas a los ojos. La Figura 5.3 muestra de manera gráfica los conflictos de la localización y descripción al estar presentes los accesorios mencionados.

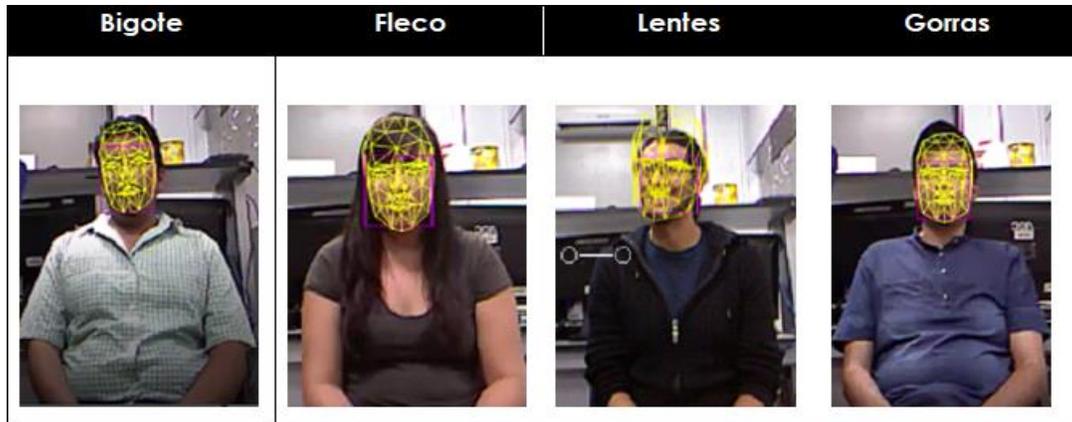


Figura 5.3 Respuesta de la máscara candide-3 ante la presencia de accesorios.

Por lo que no se considera satisfactorio utilizar el sensor Kinect con alguno de estos accesorios, ya que alteran los resultados.

5.2.2 Prueba 2: Rango de movimiento de tobillo en personas sanas

Para realizar esta prueba se utilizó el software Kinovea que es una herramienta de análisis de video dedicado al deporte, la cual sirve para analizar, hacer mediciones y comparar videos [52].

Se realizaron sesiones de video de 5 personas sanas diferentes, se les solicitó realizar los movimientos de dorsiflexión-plantarflexión. El rango solicitado fue el que cada uno de ellos alcanzará sin tener molestias y el permitido por su cuerpo. Como resultado se obtuvieron rangos del movimiento de plantarflexión entre los -35° y -45° como se puede observar en la Figura 5.4.



Figura 5.4 Rango de movimiento en personas sanas.

Por lo que se utilizan los siguientes rangos de movimientos en plantarflexión para el juego serio (Tabla 5.1):

Tabla 5.1 Rango de movimiento para cada nivel del juego serio.

Nivel del JS	Rango de movimiento
Nivel 1	De -5° a -15 °
Nivel 2	De -15.1° a -20°
Nivel 3	De -20° a -30°

5.2.3 Prueba 3: Comparación de algoritmos

El objetivo de esta prueba es realizar una comparativa de tres algoritmos de reportados en el estado del arte como buenos en la clasificación de expresiones faciales. Los tres algoritmos seleccionados son máquinas de soporte vectorial, J48 y Naïve Bayes.

Criterio de evaluación

Una métrica clave es la precisión de la clasificación general. Se define como la fracción de instancias que están clasificadas correctamente (Ecuación 5.1).

$$accuracy = \frac{sum(diag)}{n} \quad (5.1)$$

Donde:

diag= número de instancias clasificadas correctamente.

n= número de instancias.

Se utilizaron modelos en las herramientas *R* [4] y *Weka* [3] de tres algoritmos (máquinas de soporte vectorial, J48 y Naïve Bayes), utilizando una base de datos con capturas de 30 personas distintas para el entrenamiento, la cual consta de 2303 instancias. En la Tabla 5.2 se muestra la comparación de la precisión de los algoritmos generados con la base de datos de entrenamiento, en la que se puede observar que J48 tiene un mejor funcionamiento.

Tabla 5.2 Comparación de algoritmos en fase de entrenamiento.

Clasificador	Weka	R
Naïve Bayes	83.5866%	83.2826%
MSV	89.6657%	93.0881%
J48	99.0881%	99.0881%

Posteriormente se realizó una prueba con la segunda base de datos de prueba con un total de 306 instancias, la Tabla 5.2 muestra la comparación de algoritmos.

Tabla 5.3 Comparación de algoritmos en fase de pruebas.

Clasificador	Weka	R
Naïve Bayes	82.0669%	83.6601%
MSV	85.7154%	90.8496%
J48	95.7447%	91.5032%

Los resultados obtenidos por el algoritmo Naïve Bayes son los más bajos, y esto se mantiene en ambas herramientas. Por su parte, las MSV en **R** tienen un resultado superior a lo comparado con *Weka*.

La comparación en la Tabla 5.2 valida que el algoritmo J48 es la mejor opción para este sistema de reconocimiento de expresiones faciales, siendo el que tuvo un mejor desempeño en ambas herramientas teniendo una precisión de 95.7447% en *Weka* y 91.5032% en **R**, por lo que se decidió utilizar este modelo para el sistema en tiempo real, aunado a que el modelo de clasificación obtenido por este algoritmo es fácil de interpretar.

La Figura 5.5 muestra el modelo de clasificación o árbol de condiciones proporcionado por j48 y el cual se integró al sistema desarrollado, para realizar posteriormente el reconocimiento de las expresiones faciales en tiempo real.

El modelo J48 realizó una poda de aquellos variables con poca importancia en la descripción de los objetos, es decir, que lleva a cabo la selección de las variables más discriminantes. En este caso, el vector de características origina se integró de 8 variables y el modelo elimino del vector a la unidad de acción 1, la cual es referente a la caída de la mandíbula, por lo cual, el vector de características final que se trabaja se integra de 7 variables.

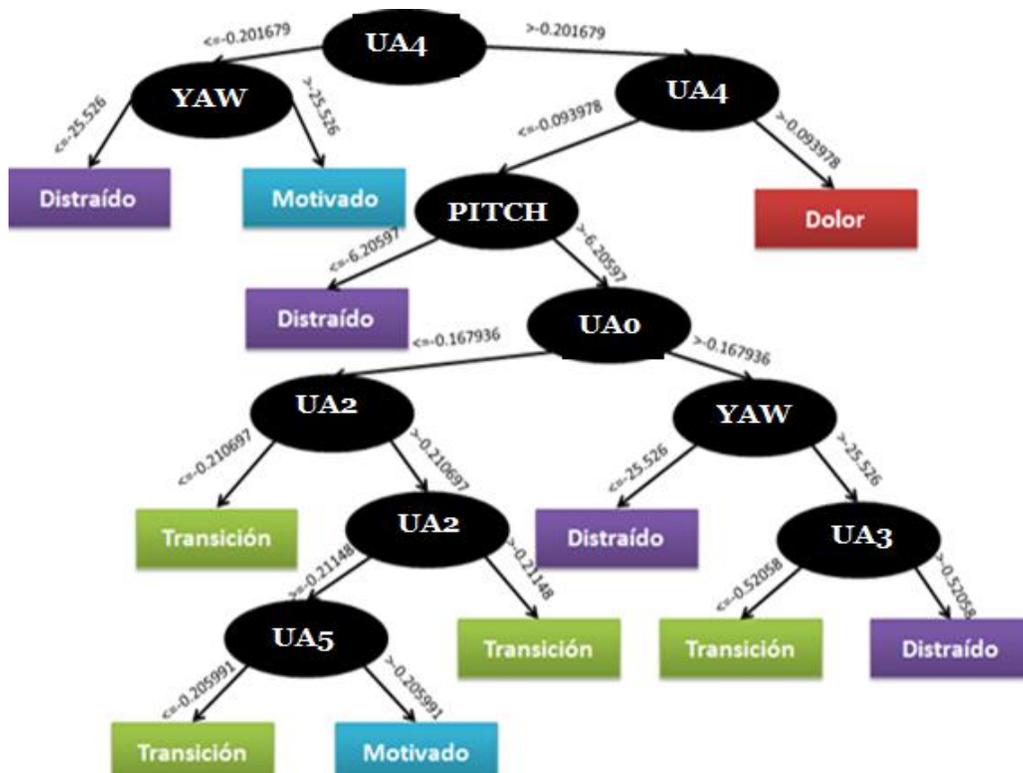


Figura 5.5 Modelo J48 generado.

5.2.4 Prueba 4: Funcionalidad de los sistemas integrados

El objetivo de esta prueba será verificar que está funcionando el sistema integral correctamente por lo que se evalúan los siguientes tres puntos:

Función 1: El objetivo fue recibir la señal de fuerza del rehabilitador la cual hace saltar al personaje principal. Se realizó la prueba de envío de datos del sensor de fuerza utilizando la herramienta que ofrece el software de *arduino* “Monitor serie” [51], mostrando el envío a *unity* haciendo saltar una figura al momento de presionar el sensor. (Figura 5.6). Los resultados fueron satisfactorios, cumpliéndose con lo especificado en el diseño del juego serio.

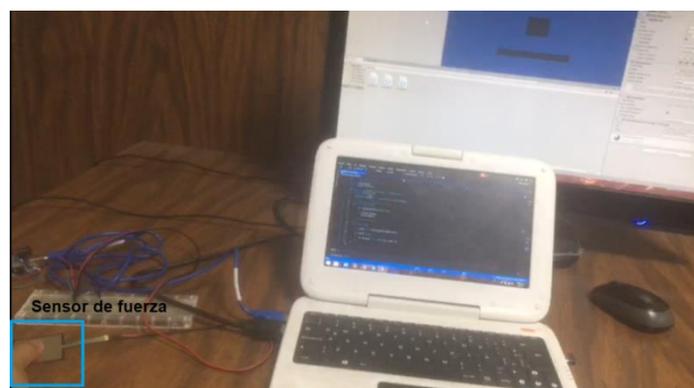


Figura 5.6 Recepción de señal de fuerza.

Función 2: El objetivo es evaluar la comunicación e integración de los tres sistemas: el sistema de reconocimiento de expresiones faciales, el juego serio y el rehabilitador de tobillo. Los sistemas de reconocimiento de expresiones faciales y el juego serio fueron integrados mediante la generación de un archivo generado en tiempo real. Es decir, el paciente inicia con su rehabilitación y a jugar. Cada segundo, el sistema de reconocimiento de expresiones identificará al menos 5 veces la expresión y tendrá un veredicto sobre si continuar con la misma frecuencia de movimientos, los aumenta o disminuye. Este archivo puede ser utilizado por el terapeuta para determinar el tiempo que duro la rehabilitación, y si el paciente está evolucionando bien o tiene alguna molestia. Por otra parte, la integración entre el juego serio y el rehabilitador de tobillo se lleva a cabo mediante un *arduino*. En la Figura 5.7 se muestran imágenes de los sistemas ya integrados.

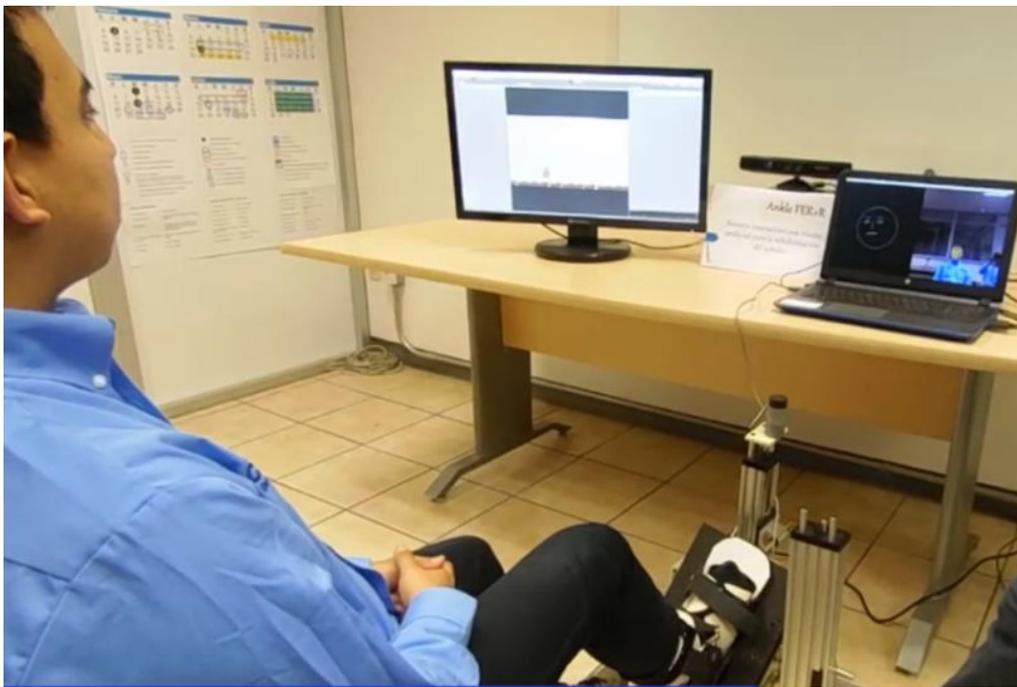


Figura 5.7 Pruebas con el sistema integrado.

Función 3: Los datos que se van almacenando en *MySQL* retroalimentan al juego serio para aumentar o disminuir la frecuencia de los obstáculos. Para lo cual se hace un promediado de los datos almacenados en 6 segundos que es la frecuencia en que aparece un obstáculo con respecto a otro, teniendo tres posibilidades:

1. **Motivado:** si el paciente se encuentra con una expresión neutral el tiempo en que aparece un obstáculo permanece igual, en el caso del nivel 1 igual a 6 segundos y de 3 segundos en el nivel 2 y 3 (Figura 5.8).

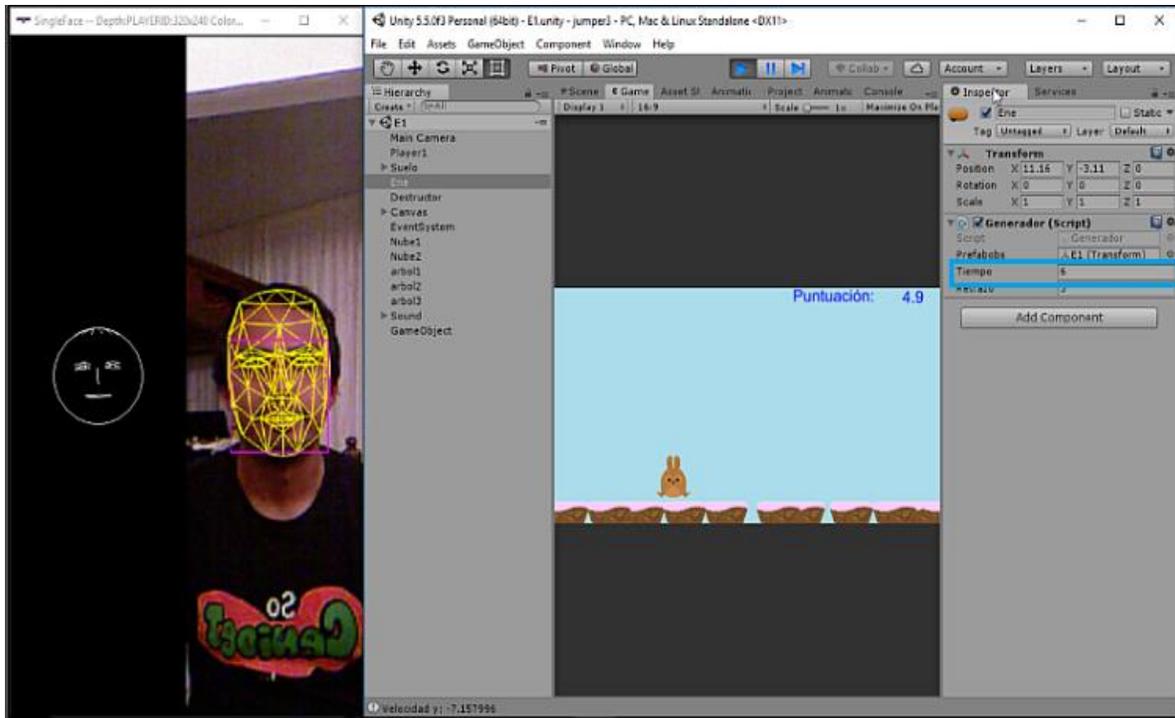


Figura 5.8 Paciente motivado en el nivel 1.

2. Desmotivado: si el paciente se encuentra distraído, el obstáculo aparece cada 3 segundos en el caso del nivel 1, en el nivel 2 y 3 su frecuencia de aparición es de cada 2 segundos (Figura 5.9).

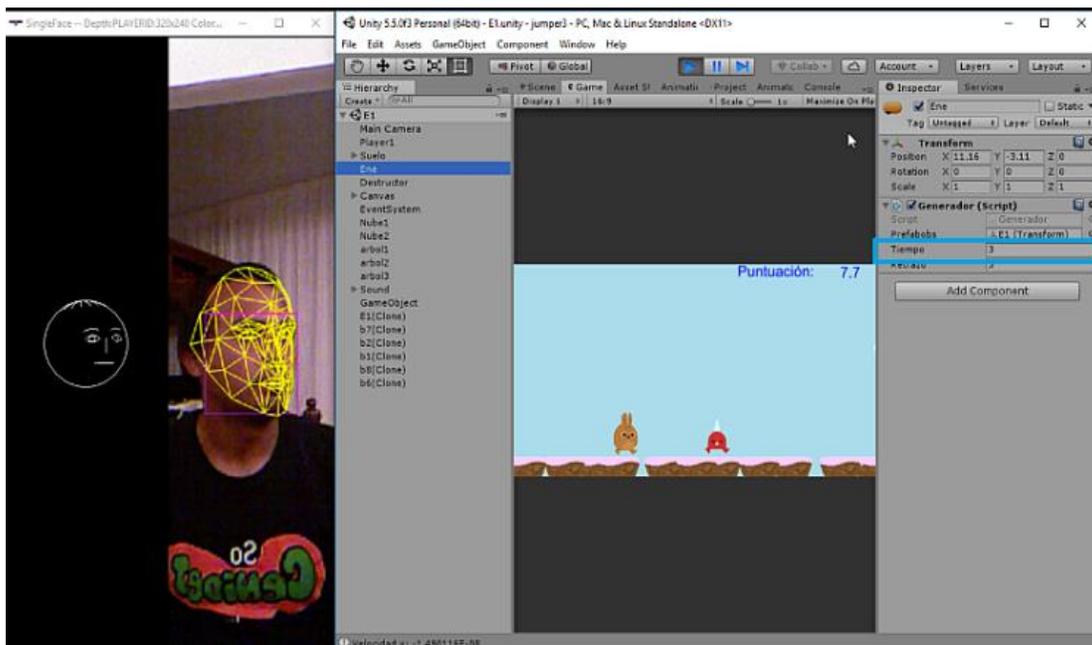


Figura 5.9 Paciente desmotivado nivel 1.

3. Dolor: si se registra que el paciente está sintiendo dolor, el obstáculo aparece cada 10 segundos y cada 6 segundos en el nivel 2 y 3 (Figura 5.10).

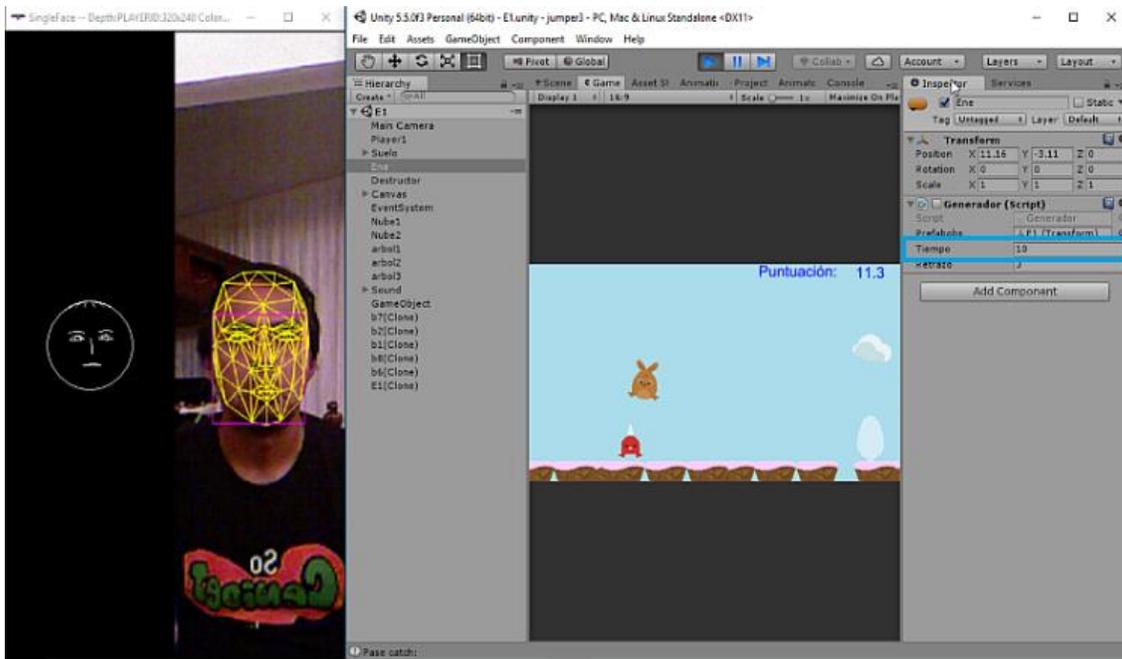


Figura 5.10 Paciente adolorido en nivel 1.

5.2.5 Prueba 5: Aplicar modelo de clasificación con los nuevos datos

El objetivo de esta prueba fue validar que el modelo J48 tiene una buena precisión para el reconocimiento de las expresiones faciales propuestas, por lo que se volvió a realizar la adquisición de video de 15 personas diferentes, en tiempo real, con el sistema integrado y comparando los nuevos datos obtenidos con el modelo creado, anteriormente, en la herramienta **R**.

Utilizando esta nueva base de datos se obtuvo una precisión de 95.6521% en J48 con el modelo creado utilizando los datos de entrenamiento por lo que este modelo es preciso para este sistema de reconocimiento.

La expresión que tiene un mejor desempeño es la de dolor siendo considerada la más importante ya que si no es detectada podría causar que el paciente se desmotive y no concluya su rehabilitación. El principal problema es la confusión con la transición, esto porque en esa categoría entran todos los movimientos de los músculos del rostro, que pueden ser expresiones no marcadas o claras, como para pertenecer a alguna de las clases propuestas.

5.3 Discusión

Este capítulo muestra las pruebas realizadas a los 3 diferentes sistemas.

Como primera prueba se realizaron estudios para visualizar las condiciones de uso de Kinect, en las cuales se determinó que debe de utilizarse con una luz artificial para no crear ningún tipo de sombra, la persona frente a Kinect tiene que estar a una distancia desde 60 cm. hasta 120 cm. y para tener un correcto funcionamiento no se debe de utilizar ningún tipo de accesorios (bigote, barba, lentes o gorro).

En la prueba 2 se realizaron sesiones de video a 5 personas diferentes para ver el rango de movimiento del tobillo en personas sanas teniendo como resultado rangos entre los -35° y los -45° .

Para la prueba 3 se utilizaron las bases de datos creadas para el entrenamiento de los algoritmos de clasificación Naïve Bayes, MSV y J48, a partir del cual se crearon tres modelos de clasificación utilizados con la base de datos de prueba, teniendo la mayor precisión con el algoritmo de J48 obteniendo un 91.5032% de precisión.

En la prueba 4 se validó la funcionalidad de los sistemas integrados, probando que el juego serio recibiera la señal de salto mediante la tarjeta *arduino* Mega y el rehabilitador. También se verificó que el sistema de visión artificial retroalimentara al juego serio, moviendo la frecuencia en que pasa un obstáculo respecto a otro mediante las expresiones faciales proporcionadas por el paciente.

Por último, en la prueba 5 se validó que J48 tiene una buena precisión para este sistema, por lo que se realizaron pruebas a 15 personas diferentes con los sistemas ya integrados y en tiempo real, obteniendo una precisión de 95.6521%.

Capítulo 6: Conclusiones

En este capítulo se presentan los objetivos alcanzados, aportaciones, las conclusiones finales y los trabajos futuros del presente trabajo.

6. 1 Objetivos alcanzados

A continuación se presenta el objetivo general de esta tesis, el cual se presentó al inicio de la investigación y su cumplimiento se mostró en las secciones de análisis, diseño e implementación de la solución, experimentación y resultados.

El objetivo principal de esta tesis fue: “Desarrollar un sistema inteligente de interacción basado en un juego serio y un sistema de visión artificial para la detección de expresiones faciales del paciente, que mejore el proceso de rehabilitación de tobillo, contemplando el movimiento de dorsiflexión-plantarflexión”.

Dicho objetivo se cumplió en su totalidad al presentar las pruebas de funcionalidad del sistema integral el cual consta de un juego serio para rehabilitación controlado por un rehabilitador de tobillo y retroalimentado por un sistema de visión artificial que a través de información no verbal del paciente, como lo son las expresiones faciales, permite saber si el paciente está sintiendo alguna molestia, contribuyendo a que concluya su rehabilitación hasta su recuperación completa.

En la Tabla 6.1 se muestran los objetivos específicos de este trabajo, así como los resultados que se obtuvieron en cada uno de ellos.

Tabla 6.1 Resultados de los objetivos específicos.

Objetivo	Comentarios
Estudiar el rehabilitador de 1 gdl.	Se realizó el estudio con el prototipo en físico el cual permite realizar el movimiento de dorsiflexión-plantarflexión de una manera activa (el paciente ejerce la fuerza). El rehabilitador es controlado por medio de un <i>arduino Mega</i> .
Revisar el estado del arte en las temáticas de reconocimiento de expresiones faciales y juegos serios para rehabilitación.	Se hizo un estudio de las dos temáticas y a partir de ellas se decidió utilizar el sensor Kinect con la librería <i>facetracking</i> , ya que ha sido utilizado en los últimos años obteniendo buenos resultados.
Analizar las expresiones faciales considerando las distintas unidades de acción	Se determinaron tres expresiones diferentes a las básicas: Motivado, Desmotivado (distráido) y con dolor.
Estudiar el <i>sdk</i> de Kinect para el reconocimiento de expresiones faciales.	Se realizó el estudio de las unidades de acción, movimientos de cabeza y condiciones para el uso de Kinect. A través de las unidades de acción identificadas con Kinect se realizó la clasificación en las tres diferentes expresiones.
Seleccionar el software de desarrollo.	Se seleccionó <i>Unity engine</i> para el desarrollo del juego serio, al utilizar Kinect <i>sdk</i> es necesario utilizar <i>visual c++</i> , también se utilizó <i>arduino</i> para realizar el controlador del rehabilitador que proporcionara señales al juego serio.
Diseñar e implementar el juego serio para rehabilitación.	Se implementó el juego serio con tres diferentes niveles de dificultad, recuperando en el nivel 1 la fuerza del tobillo, en el nivel 2 la movilidad y el nivel 3 es para reforzar la movilidad y la fuerza.
Diseñar e implementar el sistema que identifique las expresiones faciales especificadas.	Se crearon dos bases de datos una para entrenamiento y otra para pruebas y se analizaron con 3 algoritmos de clasificación (MSV, J48 y Naïve Bayes) obteniendo la mejor precisión con J48 e implementando ese modelo al sistema

6.2 Aportaciones

Se realizó un sistema inteligente para rehabilitación del tobillo, el cual es retroalimentado mediante un sistema de visión artificial, el cual indica el estado de ánimo en el que se encuentra el paciente, contribuyendo a que pueda concluir su

rehabilitación hasta su recuperación, además de que el sistema integral puede ser utilizado tanto en la clínica como en el hogar.

Los sistemas de reconocimiento de expresiones faciales revisados en el estado del arte proponen sistemas de clasificación para clasificar las expresiones faciales básicas, en este trabajo se proponen 3 expresiones faciales diferentes a las básicas.

El sistema de visión como realimentación al video juego, es una opción más económica para lograr la interacción, en comparación al uso de otros sensores como los mioeléctricos.

6.3 Trabajos futuros

En esta sección se muestran los trabajos que se pueden realizar posteriormente a este trabajo:

- Ampliar el alcance del juego serio, incluyendo los movimientos de abducción-aducción e inversión-eversión del tobillo.
- Sistema de visión artificial que detecte automáticamente el ángulo de movimiento del tobillo.
- Interfaz para que el fisioterapeuta pueda monitorear el desempeño del paciente durante sus sesiones de terapia.

6.4 Productos adicionales

- Ponencia “Desarrollo de un sistema interactivo basado en visión artificial para la rehabilitación del tobillo” en Escuela de robótica e inteligencia artificial, UTEZ, Octubre del 2017.
- Participación en el Evento Nacional Estudiantil de Innovación Tecnológica (ENEIT) con el proyecto ANKLE FER+R: Sistema interactivo con visión artificial para rehabilitación del tobillo, en la categoría producto y servicio, etapa nacional en Febrero 2018.
- Artículo **Mechatronic integral ankle rehabilitation system: ankle rehabilitation robot, serious game and facial expression recognition system** para el libro “*Advanced Topics on Computer Vision, Control and Robotic Mechatronics*”, el cual será publicado este año (2018) por Springer.

6.5 Conclusiones

- Se mejora el proceso de rehabilitación activa del tobillo, motivando al paciente mediante un juego serio.
- Se demuestra que es factible el desarrollo de un juego serio para la rehabilitación del tobillo.

- A través de la librería *facetracking* de Kinect se puede realizar el seguimiento en tiempo real de los movimientos de las partes del rostro (UAs), así como los movimientos de la cabeza, para determinar la expresión facial que está realizando el paciente.
- Se crearon y analizaron tres bases de datos, las cuales fueron analizadas con tres diferentes algoritmos de clasificación (MSV, Naïve Bayes y J48) teniendo cuatro diferentes clases: motivado, desmotivado, con dolor y transiciones.
- Se mostró que para este sistema de reconocimiento de expresiones faciales el mejor resultado fue utilizando J48 con una precisión del 95.6521% para las tres expresiones faciales especificadas con el sistema integrado.
- El juego serio es retroalimentado con la expresión facial, la cual dependiendo de lo que el paciente esté expresando indica la frecuencia en que aparece un obstáculo respecto al otro.

Referencias

- [1] F. Gómez, “Diseño y Construcción de un Mecanismo para la Rehabilitación de Tobillo Basado en una Mesa X-Y,” 2013.
- [2] C. Guzmán, A. Blanco, M. Oliver, J. Carrera, and F. Gómez, “Diseño y control de un sistema interactivo para la rehabilitación de tobillo: TobiBot Design,” *Ing. Mecánica, Tecnol. y Desarro.*, vol. 5(1), pp. 255–264, 2014.
- [3] “Weka.” [Online]. Available: <https://www.cs.waikato.ac.nz/ml/weka/>.
- [4] “r studio.” [Online]. Available: <https://www.rstudio.com/>.
- [5] A. Porras, “Sistema de reconocimiento de expresiones faciales aplicado a la interacción humano-computadora usando redes neuronales y flujo óptico,” UNAM, 2005.
- [6] B. Fasel and J. Luetttin, “Automatic facial expression analysis : a survey,” vol. 36, pp. 259–275, 2003.
- [7] P. Michel and R. El Kaliouby, “Real time facial expression recognition in video using support vector machines,” *Proc. 5th Int. Conf. Multimodal interfaces - ICMI '03*, pp. 258–264, 2003.
- [8] F. Tsalakanidou and S. Malassiotis, “Real-time 2D + 3D facial action and expression recognition,” *Pattern Recognit.*, vol. 43, no. 5, pp. 1763–1775, 2010.
- [9] S. Gupta, K. Verma, and N. Perveen, “Facial Expression Recognition System Using Facial Characteristic Points And ID3,” no. 1, pp. 45–49, 2012.
- [10] M. Lyons, M. Kamachi, J. Gyoba, and R. Kubota, “The Japanese Female Facial Expression (JAFFE) Database Publications with information about the JAFFE database How to obtain access to the images : Related Research Projects on this Site ;,” pp. 1–2, 2016.
- [11] V. Surbhi, “Roi Segmentation for Feature Extraction,” *arXiv Prepr. arXiv1207.2922.*, 2012.
- [12] Á. A. Arenas, B. J. Cotacio, E. S. Isaza, J. V. Garcia, J. A. Morales, and J. I. Marín, “Sistema de Reconocimiento de Rostros en 3D usando Kinect,” *Symp. IMAGE, SIGNAL Process. Artif. Vis.*, vol. XVII, 2012.
- [13] B. Seddik, H. Maamatou, S. Gazzah, T. Chateau, and N. E. Ben Amara, “Unsupervised facial expressions recognition and avatar reconstruction from Kinect,” in *Systems, Signals & Devices (SSD), 2013 10th International Multi-Conference on*, 2013, pp. 1–6.
- [14] D. Li, C. Sun, F. Hu, D. Zang, L. Wang, and M. Zhang, “Real-time

- performance-driven facial animation with 3ds Max and Kinect,” *2013 3rd Int. Conf. Consum. Electron. Commun. Networks, CECNet 2013 - Proc.*, pp. 473–476, 2013.
- [15] E. P. Ijjina and C. K. Mohan, “Facial Expression Recognition Using Kinect Depth Sensor and Convolutional Neural Networks,” in *Machine Learning and Applications (ICMLA), 2014 13th International Conference*, 2014, pp. 392–396.
- [16] L. Stocchi, “3D Facial Expressions Recognition Using the Microsoft Kinect,” *Sch. Comput. Dublin City Univ.*, vol. Artículo n, 2014.
- [17] A. Tcherkassof, D. Dupré, B. Meillon, N. Mandran, M. Dubois, and J.-M. Adam, “Dynemo : A Video Database of Natural Facial Expressions of Emotions,” *Int. J. Multimed. Its Appl.*, vol. 5, no. 5, pp. 61–80, 2013.
- [18] C. Cao, Y. Weng, S. Zhou, Y. Tong, and K. Zhou, “FaceWarehouse: A 3D facial expression database for visual computing,” *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 3, pp. 413–425, 2014.
- [19] S. Aly, A. Youssef, and L. Abbott, “ADAPTIVE FEATURE SELECTION AND DATA PRUNING FOR 3D FACIAL EXPRESSION RECOGNITION USING THE KINECT Bradley Department of Electrical and Computer Engineering , Virginia Tech , Blacksburg , Virginia , USA The City of Scientific Research and Technological App,” pp. 1361–1365, 2014.
- [20] M. Kakarla and G. R. M. Reddy, “A real time facial emotion recognition using depth sensor and interfacing with Second Life based Virtual 3D avatar,” *Int. Conf. Recent Adv. Innov. Eng.*, pp. 1–7, 2014.
- [21] S. Aly, A. Trubanova, L. Abbott, S. White, and A. Youssef, “VT-KFER: A Kinect-based RGBD+ Time dataset for spontaneous and non-spontaneous facial expression recognition,” in *2015 International Conference on Biometrics (ICB)*, 2015, pp. 90–97.
- [22] Q. Mao, X. Pan, Y. Zhan, and X. Shen, “Using Kinect for real-time emotion recognition via facial expressions,” *Front. Inf. Technol. Electron. Eng.*, vol. 16, no. 4, pp. 272–282, 2015.
- [23] J. Kumari, R. Rajesh, and K. M. Pooja, “Facial Expression Recognition: A Survey,” *Procedia Comput. Sci.*, vol. 58, pp. 486–491, 2015.
- [24] T. R. Patil and S. S. Sherekar, “Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification,” vol. 6, no. 2, 2013.
- [25] E. Carmona Suarez, “Tutorial sobre Máquinas de Vectores Soporte (SVM),” *Dpto. Intel. Artif. ETS Ing. Informática, Univ. Nac. Educ. a Distancia*, 2014.
- [26] C. Hsu, C. Chang, and C. Lin, “A Practical Guide to Support Vector Classification,” vol. 1, no. 1, pp. 1–16, 2010.

- [27] N. Bhargava and G. Sharma, “Decision Tree Analysis on J48 Algorithm for Data Mining,” *Int. J. Adv. Res.*, vol. 3, no. 6, pp. 1114–1119, 2013.
- [28] J. S. Santana and E. M. Farfán, “El Arte de programar en R,” p. 182, 2014.
- [29] K. P. Michmizos and H. I. Krebs, “Serious games for the pediatric anklebot,” *Proc. IEEE RAS EMBS Int. Conf. Biomed. Robot. Biomechatronics*, pp. 1710–1714, 2012.
- [30] L. Omelina, B. Jansen, B. Bonnechère, S. V. S. Jan, and J. Cornelis, “Serious games for physical rehabilitation: designing highly configurable and adaptable games,” pp. 10–12, 2012.
- [31] G. C. Burdea, S. Member, D. Cioi, A. Kale, W. E. Janes, S. A. Ross, and J. R. Engsborg, “Robotics and Gaming to Improve Ankle Strength , Motor Control , and Function in Children With Cerebral Palsy — A Case Study Series,” vol. 21, no. 2, pp. 165–173, 2013.
- [32] J. Bertrand, L. C. Dukes, P. Dukes, E. Ebrahimi, A. Hayes, N. Mack, J. McClendon, D. Parmar, T. B. Pence, B. Shannon, A. Wachter, Y. Wu, S. Babu, and L. F. Hodges, “Serious Games for Training , Rehabilitation and Workforce Development,” *IEEE VIRTUAL Real.*, pp. 195–196, 2013.
- [33] E. Vogiatzaki and A. Krukowski, “Serious Games for Stroke Rehabilitation Employing Immersive User Interfaces in 3D Virtual Environment,” *J. Heal. Informatics*, vol. 6, pp. 1–8, 2013.
- [34] A. C. B. F. Gonçalves, S. Carlos, W. M. Santos, L. J. Consoni, and A. A. G. Siqueira, “Serious Games for Assessment and Rehabilitation of Ankle Movements,” *Serious Games Appl. Heal.*, no. IEEE 3rd International Conference, pp. 1–6, 2014.
- [35] J. A. Garcia, “The Mobile RehApp™: an AR-based mobile game for ankle sprain rehabilitation,” *Serious Games Appl. Heal.*, no. IEEE 3rd International Conference, pp. 1–6, 2014.
- [36] M. Zhang, G. Zhu, A. Nandakumar, S. Gong, and S. Xie, “A virtual-reality tracking game for use in robot-assisted ankle rehabilitation,” *MESA 2014 - 10th IEEE/ASME Int. Conf. Mechatron. Embed. Syst. Appl. Conf. Proc.*, 2014.
- [37] N. Shah, “Designing Motivational Games for Stroke Rehabilitation,” *7th Int. Conf. Hum. Syst. Interact.*, pp. 166–171, 2014.
- [38] A. B. Farjadian, M. Nabian, M. K. Holden, and C. Mavroidis, “Development of 2-DOF Ankle Rehabilitation System,” *Annu. Northeast Bioeng. Conf.*, vol. 40th, no. IEEE, pp. 2–3, 2014.
- [39] R. C. Menezes, P. K. A. Batista, A. Q. Ramos, and A. F. C. Medeiros, “Development of a complete game based system for physical therapy with

- kinect,” *SeGAH 2014 - IEEE 3rd Int. Conf. Serious Games Appl. Heal. Books Proc.*, pp. 1–6, 2014.
- [40] A. Jaume-i-capó and A. Sam, “Vision-based interaction as an input of serious game for motor rehabilitation,” *Telecommun. Forum Telfor*, vol. 22nd, no. IEEE, pp. 854–857, 2014.
- [41] H. Tannous, T. T. Dao, D. Istrate, and M. H. B. Tho, “Serious Game for Functional Rehabilitation,” *2015 Int. Conf. Adv. Biomed. Eng.*, pp. 242–245, 2015.
- [42] A. M. Agas, A. A. Daitol, U. B. Shah, L. J. Fraser, K. M. Abbruzzese, K. K. Karunakaran, and R. Foulds, “3-DOF Admittance Control Robotic Arm with a 3D Virtual Game for facilitated training of the hemiparetic hand,” *Annu. Northeast Biomed. Eng. Conf.*, vol. 41st, no. IEEE, pp. 1–2, 2015.
- [43] T. B. Pasqual, G. A. P. Caurin, and A. A. G. Siqueira, “Serious Game Development for Ankle Rehabilitation Aiming at User Experience,” *Int. Conf. Biomed. Robot. Biomechatronics*, vol. 6th, pp. 1015–1020, 2016.
- [44] M. J. Girone, G. C. Burdea, and M. Bouzit, “The ‘rutgers ankle’ orthopedic rehabilitation interface,” *Stud. Health Technol. Inform.*, vol. 3, 1999.
- [45] V. M. Penagos, “Diseño y construcción de un robot paralelo con dos guías lineales,” Cenidet, 2016.
- [46] J. Hervas, “Psicología del color en videojuegos,” *Equipo Illusion Studio*, 2017. [Online]. Available: www.psicologiadelcolor.es/articulos/psicologia-del-color-en-videojuegos.
- [47] Microsoft, “Face Tracking,” 2017. [Online]. Available: [https://msdn.microsoft.com/en-us/library/jj130970\(d=printer\).aspx](https://msdn.microsoft.com/en-us/library/jj130970(d=printer).aspx).
- [48] Microsoft, “Visual C++,” 2015. [Online]. Available: <https://www.microsoft.com/es-mx/download/details.aspx?id=48145>.
- [49] F. Webinars, “MySQL,” 2017. [Online]. Available: <https://www.mysql.com/>.
- [50] Unity, “Unity Engine,” 2017. [Online]. Available: <https://connect.unity.com/>.
- [51] Arduino, “Arduino the online IDE,” 2017. [Online]. Available: <https://www.arduino.cc>. [Accessed: 01-Jan-2017].
- [52] Kinovea, “Kinovea,” 2017. [Online]. Available: <https://www.kinovea.org/>.
- [53] I. Mpiperis, S. Malassiotis, and M. G. Strintzis, “Bilinear Models for 3-D Face and Facial Expression Recognition,” *IEEE Trans. Inf. Forensics Secur.*, vol. 3, no. 3, pp. 498–511, 2008.
- [54] A. Dhall, R. Goecke, J. Joshi, M. Wagner, and T. Gedeon, “Emotion recognition in the wild challenge 2013,” *Int. Conf. Multimodal Interact.*, pp.

509–516, 2013.

- [55] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, “The extended cohn-kande dataset (CK+): A complete facial expression dataset for action unit and emotionspecified expression,” *IEEE Conf. Comput. Vis. Pattern Recognit. Work.*, no. July, pp. 94–101, 2010.
- [56] H. Search, C. About, C. Request, T. Mmi, F. Expression, A. Human, B. Analysis, M. Expression, M. M. I. Facial, E. Database, F. A. Unit, and A. Descriptors, “MMI Facial Expression Database,” <http://mmifacedb.eu/>, 2016. .
- [57] Queens University Belfast, “Belfast Naturalistic,” <http://belfast-naturalistic-db.sspnet.eu/>, 2016. .
- [58] A. Aifanti, N., Papachristou, C., & Delopoulos, “The MUG Facial Expression Database,” *Image Anal. Multimed. Interact. Serv.*, no. May 2010, pp. 1–4, 2010.
- [59] B. La Rochelle, “3D Twins Expression Challenge (3D - TEC) 2 . Affectiva - MIT Facial Expression Dataset (AM - FED) (2013) 3 . A Natural Visible and Infrared facial Expression Database (2010) 6 . Binghamton University Facial Expression Databases : BP4D - Spontaneous,” no. 2010, pp. 1–11, 2016.
- [60] D. McDuff, R. El Kaliouby, T. Senechal, M. Amr, J. F. Cohn, and R. Picard, “Affectiva-mit facial expression dataset (AM-FED): Naturalistic and spontaneous facial expressions collected ‘in-the-wild,’” *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, pp. 881–888, 2013.
- [61] Riumn, “EURECOM Kinect Face Dataset,” <http://rgb-d.eurecom.fr/>, 2016. [Online]. Available: <http://rgb-d.eurecom.fr/>.
- [62] S. Shokrani, P. Moallem, and M. Habibi, “Facial emotion recognition method based on Pyramid Histogram of Oriented Gradient over three direction of head,” *Proc. 4th Int. Conf. Comput. Knowl. Eng. ICCKE 2014*, pp. 215–220, 2014.
- [63] V. LoBue and C. Thrasher, “The Child Affective Facial Expression (CAFE) set: Validity and reliability from untrained adults,” *Front. Psychol.*, vol. 6, no. JAN, pp. 1–9, 2015.
- [64] S. Setty and et al., “Indian Movie Face Database: A Benchmark for Face Recognition Under Wide Variations,” *Ncvprimg*, no. i, 2013.

Anexos

ANEXO I: Bases de datos para reconocimiento de expresiones faciales

Nombre	Desarrollada por:	E/D	2D/3D	# de imágenes o videos	Modelos	Edad	Expresiones	Información adicional
BU-3DF [53]	Binghamon University	Estadística	3D	500	100	18-70	Enojo, disgusto, temor, felicidad, tristeza, sorpresa	Participantes con distintas etnias.
BU-4DF [53]	Binghamon University	Dinámica	3D+tiempo	606	101	N/A	Enojo, disgusto, temor, felicidad, tristeza, sorpresa	Secuencias de imágenes de personas de diferente sexo y etnias.
BP4D-Spontaneous [53]	Binghamon University	N/A	3D/2D	N/A	41	18-29	8 emociones	Base de datos en video (2.6 TB), con personas de diferentes etnias (23 mujeres y 18 hombres).
EmotiW [54]	Abhinav Dhal	Estática	2D	518 entrenamiento y 383 validación	N/A	N/A	Enojo, disgusto, temor, felicidad, tristeza, sorpresa	Imágenes con varias posiciones del rostro.
Cohn-Kanade [55]	Jeff Cohn, Carnegie Mellon University	N/A	2D	486	100	N/A	23 unidades de acción (solas o combinadas)	486 videos, tomas de 12 cuadros por segundo, los modelos son universitarios (65% mujeres, 15% africanos, 3% latinos y asiáticos)
MMI [56]	M. Pantic, F. Valster	Estática	2D	1280 videos, 250 imágenes	43	N/A	79 unidades de acción	24 cuadros por segundo, con vista frontal y de perfil.
JAFFE [10]	Miyuki Kamachi, Michael Lyons, Jiro Gyoba	Estática	2D	213	10	N/A	Enojo, disgusto, temor, felicidad, tristeza, sorpresa	Solo incluye fotos de mujeres con vista frontal.
Belfast Naturalistic [57]	Queens University Belfast	Estática	2D	>250	125	N/A	Una amplia gama de expresiones	Fotos espontaneas de expresiones de 31 hombres y 94 mujeres.
MUG Facial Expression [58]	Multimedia Understanding Group	N/A	2D	>1462	52	20-25	F Enojo, disgusto, temor, felicidad, tristeza, sorpresa	Fotos de 35 mujeres y 51 hombres.
3D-Tec [59]	The University of Notre Dame du Lac	N/A	3D	428	214	N/A	Felicidad, neutral	Fotos de 107 pares de gemelos espontaneas.
AM-FED [60]	Daniel McDuff	Estática	2D	242	N/A	N/A	10 FACS simétricas, 4 asimétricas, 2 movimientos de cabeza	Capturadas espontáneamente durante los modelos veían un video.
DynEmo [17]	Tchekassof, Meillon, Mandran	Dinámica	2D	358	358	25-65	Enojo, disgusto, temor, felicidad, tristeza, sorpresa	182 mujeres y 176 hombres caucásicos.
EURECOM [61]	Riurnn	N/A	3D	N/A	52	N/A	Felicidad, neutral, boca abierta y con oclusiones en ojos y boca	Capturada con sensor Kinect secuencias de 14 mujeres y 38 hombres.
RaFD [62]	The Radboud University Nijmegen	Dinámica	2D	N/A	67	21 promedio	Enojo, disgusto, temor, felicidad, tristeza, sorpresa	Modelos entre niños, mujeres y hombres caucásicos.
CAFÉ [63]	Vanessa LoBue	N/A	2D	1200	100	2-8	Enojo, disgusto, temor, felicidad, tristeza, sorpresa	Modelos niños utilizada para búsqueda científica
IMFDB [64]	Shankar Setty	N/A	2D	34512	100	N/A	Enojo, disgusto, temor, felicidad, tristeza, sorpresa	Imágenes con posición, escala, expresión e iluminación variada.