

UNIVERSIDAD ANDRÉS BELLO

FACULTAD DE INGENIERÍA

INGENIERIA EN COMPUTACIÓN E INFORMÁTICA

**Sistema inteligente para el apoyo en el diagnóstico en exámenes de Inmunofluorescencia para Anticuerpos Autoinmunes**

**PROYECTO DE TÍTULO PARA OPTAR AL TÍTULO DE**

**Ingeniero en Computación e Informática**

Autor:

César Germán Cheuque Cerda

Profesor Guía:

Ing. Romina Torres

VIÑA DEL MAR, 2017

Agradecimientos

Dedicatoria

**Índice de Contenidos.**

Contenido

[Resumen 10](#_Toc502335968)

[Capítulo I 12](#_Toc502335969)

[1.1 Introducción. 13](#_Toc502335970)

[Capitulo II: Fundamentación del Tema 15](#_Toc502335971)

[2.1. Contexto 16](#_Toc502335972)

[2.1.1 Fundamentos de Laboratorio 18](#_Toc502335973)

[2.1.2 Fundamentos de Inmunología 19](#_Toc502335974)

[2.1.3. Enfermedades Autoinmunes. 25](#_Toc502335975)

[2.2. Reconocimiento de imágenes. 26](#_Toc502335976)

[2.3. Reconocimiento del problema, necesidad y oportunidad. 30](#_Toc502335977)

[2.3.1 Posibles soluciones 33](#_Toc502335978)

[2.3.2. Alternativas de solución ideal. 34](#_Toc502335979)

[2.3.3 Solución Propuesta 35](#_Toc502335980)

[2.4. Objetivos 36](#_Toc502335981)

[2.5.- Alcance: 37](#_Toc502335982)

[2.6. Limitaciones 37](#_Toc502335983)

[2.7. Requerimientos 38](#_Toc502335984)

[Capitulo III: Materiales y Métodos 40](#_Toc502335985)

[3.1 Selección de metodologías 41](#_Toc502335986)

[3.2 Selección de técnicas y tecnologías. 41](#_Toc502335987)

[3.3 Matriz de trazabilidad de requerimientos 42](#_Toc502335988)

[3.4 Planificación del proyecto 43](#_Toc502335989)

[3.5 Métricas de los objetivos. 44](#_Toc502335990)

[3.6 Plan de Gestión de la Configuración 46](#_Toc502335991)

[3.7 Plan Gestión de Cambios. 47](#_Toc502335992)

[3.8 Plan de Gestión de Riesgos 48](#_Toc502335993)

[3.9 Ambiente de desarrollo, pruebas y producción 51](#_Toc502335994)

[Capitulo IV: Desarrollo y Resultados 52](#_Toc502335995)

[4.1 Desarrollo de los Sprint. 53](#_Toc502335996)

[Capítulo V: Discusiones. 78](#_Toc502335997)

[5.1 Discusiones. 79](#_Toc502335998)

[Capítulo VI: Conclusiones 84](#_Toc502335999)

[6.1 Conclusiones y trabajos posteriores 85](#_Toc502336000)

[Bibliografía. 87](#_Toc502336001)

**Índice de Figuras**

[Figura 1. Esquema de la estructura básica de una Inmunoglobulina (Modificado de Koivunen M., y Krogsrud R., 2006) 20](#_Toc502335910)

[Figura 2. Gráfico comparativo de 5 fluorocromos distintos bajo una luz de 480nm. Realizado con la herramienta online de BD Biosciences. 22](#_Toc502335911)

[Figura 3. Ejemplo de IFD (I) e IFI (II). (Koivunen M., y Krogsrud R., 2006) 23](#_Toc502335912)

[Figura 4 Se muestra un examen de Inmunofluorescencia Directa (IFD), en el cual las células que brillan de color verde indican la presencia del Virus Respiratorio Sincicial (Vircell Microbiologists) 24](#_Toc502335913)

[Figura 5. Patrones de unión en Inmunofluorescencia Indirecta (IFI). (Koivunen M., y Krogsrud R., 2006) 25](#_Toc502335914)

[Figura 6. Número de papers sobre Deep Learning por técnica utilizada, realizados entre 2012 y Junio de 2017. (Litjens, G. y cols., 2017) 29](#_Toc502335915)

[Figura 7. Ishikawa del problema. Elaboración propia 31](#_Toc502335916)

[Figura 8. Propuesta del sistema (Elaboración propia) 36](#_Toc502335917)

[Figura 9. Diagrama de Casos de Uso 55](#_Toc502335918)

[Figura 10. Diagrama de Despliegue del Sistema 59](#_Toc502335919)

[Figura 11. Modelo de Base de Datos. Elaboración propia 60](#_Toc502335920)

[Figura 12. Burndown Chart del Sprint 1. 64](#_Toc502335921)

[Figura 13. Estructura del Proyecto. 66](#_Toc502335922)

[Figura 14. Página de Login del Sistema 67](#_Toc502335923)

[Figura 15. Página principal de la Aplicación Web. 67](#_Toc502335924)

[Figura 16. Burndown Chart del Sprint 2. 69](#_Toc502335925)

[Figura 17. Representación gráfica del Modelo desarrollado para la Clasificación de los Patrones Homogéneo y Moteado. 74](#_Toc502335926)

[Figura 18. Precisión del Modelo 75](#_Toc502335927)

[Figura 19. Pérdida del Modelo 75](#_Toc502335928)

[Figura 20. Burndown Chart del Sprint 3. 77](#_Toc502335929)

[Figura 21. Evolución del Riesgo durante el Desarrollo del Proyecto 82](#_Toc502335930)

**Índice de Tablas**

[Tabla 1: Características de las inmunoglobulinas (Palomo I., y cols., 2002) 20](#_Toc502335931)

[Tabla 2. Comparación de diversas soluciones (Elaboración propia) 33](#_Toc502335932)

[Tabla 3. Comparativa de algoritmos inteligentes (Creación Propia) 34](#_Toc502335933)

[Tabla 4.Matriz de trazabilidad de requerimientos 42](#_Toc502335934)

[Tabla 5. Identificación descripción del riesgo técnico. 48](#_Toc502335935)

[Tabla 6. Planificación de medidas de mitigación del riesgo técnico 49](#_Toc502335936)

[Tabla 7. Historias de Usuario 53](#_Toc502335937)

[Tabla 8. Pila del Sprint 1, la que sigue los pasos para realizar la Historia de Usuario 5. 54](#_Toc502335938)

[Tabla 9. Pila del Sprint 2, la que sigue los pasos para realizar la Historia de Usuario 6. 65](#_Toc502335939)

[Tabla 10. Pila del Sprint 3, la que sigue los pasos para realizar la Historia de Usuario 4. 70](#_Toc502335940)

[Tabla 11. Matriz de Riesgo (Riesgo = Impacto \* Probabilidad de Ocurrencia) 81](#_Toc502335941)

**Control de Versiones.**

|  |  |  |  |
| --- | --- | --- | --- |
| Versión | Responsable | Detalle | Fecha |
| 0.1 | César Cheuque C. | Capítulo 1 y II | 30-03-17 |
| 0.2 | César Cheuque C. | Corrección de Capítulos | 15-04-17 |
| 0.3 | César Cheuque C. | Corrección O.G y O.E | 20-04-17 |
| 0.4 | César Cheuque C. | Sprint Backlog | 20-05-17 |
| 0.5 | César Cheuque C. | Capítulo III | 22-05-17 |
| 0.6 | César Cheuque C. | Capítulo IV | 20-06-17 |
| 0.7 | César Cheuque C. | Refactoring Tesis | 05-08-17 |
| 0.75 | César Cheuque C. | Refactoring Capitulo II | 10-08-17 |
| 0.77 | César Cheuque C. | Refactoring Capítulo III | 20-10-17 |
| 0.9 | César Cheuque C. | Capítulo IV | 20-11-17 |
| 1.0 | César Cheuque C. | Capítulo V y VI | 29-11-17 |

# Resumen

A lo largo de los años el uso de nuevas tecnologías y en particular el de las tecnologías de la información aplicadas, han permitido la automatización de múltiples procesos en un sinfín de áreas. Esto cobra aún mayor relevancia en lo que respecta al área de la salud, y en este caso en lo que respecta a los laboratorios, donde la mejora continua de procesos ha permitido aumentar la capacidad de procesamiento, abaratamiento de costos y reducir considerablemente la incidencia de errores.

Sin embargo, pese a lo mucho que se ha logrado avanzar en la automatización de procesos en el laboratorio, aún existen técnicas que debido a su complejidad y/o el elevado costo que implicaría el desarrollo de técnicas alternativas, ha hecho que se siga trabajando de la misma forma que se hacía hace más de 30 años, este es el caso de las pruebas de Inmunofluorescencia y las de VDRL, pese a su importancia diagnóstica y alta demanda.

En la actualidad existen exámenes que funcionan como alternativas a los ya mencionados, pero la mayoría son poco prácticos para su uso diario, además de muy costosos, tanto para el laboratorio como para los pacientes, y no representan una verdadera mejora como apoyo en el diagnóstico. Respecto al uso de tecnologías informáticas para el apoyo de técnicas de lectura visual, se han desarrollado un par de experiencias al respecto pero referidas al reconocimiento de células, pero no al de patrones de exámenes.

En este trabajo se propone la realización de un sistema de reconocimiento de patrones de positividad/reactividad de exámenes de laboratorio, en específico para la prueba de Inmunofluorescencia para Anticuerpos Autoinmunes, en base al uso de algoritmos inteligentes de auto-aprendizaje que sirvan de apoyo diagnóstico al profesional a cargo de informar estos exámenes.

Para la validación de las técnicas desarrolladas durante este trabajo, se usarán imágenes cuyo resultado será contrastado con la opinión de expertos en el área.

La contribución que este proyecto aportará está en reducir aún más la probabilidad de errores en el informe, a la vez de representar una alternativa de bajo coste a la lectura manual de estos exámenes.

# Capítulo I

## 1.1 Introducción.

Las pruebas de laboratorio se han beneficiado por el uso de las nuevas tecnologías, lo que ha permitido que se reduzca cada vez más la posibilidad de cometer errores, al mismo tiempo ha permitido reducir los costos y aumentar el volúmen de exámenes posibles de realizar, lo que en conjunto significa una mejora en la calidad y en la accesibilidad de estos a los pacientes. Sin embargo, a día de hoy aún existen exámenes que, dadas sus características, aún no se ven beneficiadas por el uso de los avances modernos, realizándose de la misma forma que lo hacía hace más de 20 años.

En este trabajo se automatizará la clasificación de los resultados de Inmunofluorescencia para Anticuerpos Autoinmunes utilizando técnicas de Inteligencia Artificial para el reconocimiento de los patrones de positividad e interpretación de estas pruebas, ya que por su importancia médica y volumen de muestras hacen de la automatización de la lectura de estos exámenes resulta relevante.

En los diferentes capítulos de esta tesis se describen los siguientes contenidos:

**Capítulo 1:** Se entrega el contexto del problema y las justificaciones del proyecto que se realizará

**Capítulo 2**: Se describirá la fundamentación del tema a abordar.

**Capítulo 3**: Se describen los materiales y métodos utilizados

**Capítulo 4**: Se Muestra el desarrollo y resultados obtenidos

**Capítulo 5**: Se Discuten resultados

**Capítulo 6**: Se realizan conclusiones y se estiman trabajos futuros

# Capitulo II: Fundamentación del Tema

## 2.1. Contexto

Los laboratorios surgieron como complemento al diagnóstico del médico, pero actualmente también tienen alta relevancia en el control de la evolución del tratamiento médico y en la prevención de las distintas enfermedades, permitiendo a los médicos tomar decisiones de forma correcta, lo que conlleva que el trabajo en los laboratorios deba ser confiable, oportuno y confidencial (Micucci, H. y cols. 2010). Los pilares de cualquier laboratorio son la tecnología que se usa y el recurso humano. En Chile son los Tecnólogos Médicos los profesionales autorizados para la realización de los exámenes, y los Técnicos en Laboratorio son quienes brindan las funciones de apoyo en las labores del Laboratorio. El trabajo de estos laboratorios se divide en secciones que usualmente son: Química Clínica, Hematología, Microbiología, Inmunohematología, Serología, Parasitología, Uroanálisis y Urgencias. Actuando como jefe del laboratorio un Director Técnico que puede se un Médico, Químico Farmacéutico, Bioquímico o Tecnólogo Médico.

En lo que respecta a la tecnología, la gran revolución de los laboratorios vino de la mano del uso de equipamientos y nuevas tecnologías que permitieron la automatización de las técnicas, permitiendo así abaratar costos, reducir los tiempos de respuesta y aumentar el volumen de exámenes posibles de realizar en el mismo periodo de tiempo, a modo de ejemplo hasta hace unos años la realización de un único hemograma implicaba varios pasos que debían realizarse metódicamente llegando a demorarse una hora exacta el paso que más tiempo llevaba. A día de hoy, es posible, dependiendo del caso, tener un hemograma en algo así de 10 a 15 minutos y procesar en una hora unos 50 a 100 hemogramas a una fracción del costo, tanto en los insumos utilizados como en las horas-hombre utilizadas.

Otra consideración a tener en cuenta es que como cualquier empresa, los recursos de los laboratorios son limitados por lo que estos deben ser usados de manera inteligente tratando de maximizar la cantidad de exámenes realizados con el mínimo costo posible.

A día de hoy existen algunas técnicas que se siguen realizando de forma manual como las reacciones de aglutinación e inmunofluorescencia, basadas en reacciones antígeno-anticuerpo, pese a que estas técnicas son ampliamente utilizadas. Los motivos que han hecho que estas pruebas no se hayan actualizado como otras, radica en primera instancia en que el reconocimiento de los patrones que indican la positividad o reactividad de las pruebas es bastante complejo, requiriendo muchas horas de capacitación por parte del Tecnólogo Médico que realizará su lectura, y por otra parte al hecho de que otras alternativas son demasiado costosas tanto en equipamiento como en los recursos utilizados, además que no entregan necesariamente la misma información al médico tratante.

### 2.1.1 Fundamentos de Laboratorio

**Conceptos básicos.**

En el laboratorio clínico y en otras ramas de la ciencia existen pruebas de distintos tipos que aún cuando sirven para pesquizar el mismo objeto de análisis tienen distintas características que las hacen idóneas para tal o cual propósito, y aunque el costo, las instalaciones y el equipamiento necesario son sumamente importantes a la hora de escoger una opción existen otras características fundamentales que deben ser consideradas, la sensibilidad y la especificidad:

* Sensibilidad: la sensibilidad de una técnica de laboratorio está definida por la capacidad de detectar un analito cuando este está presente en la muestra evaluada.
* Especificidad: la específicidad de una técnica es la capacidad de reaccionar al analito que se quiere detectar y nada más que a éste.

En el laboratorio clínico es de suma importancia que una técnica en particular sea lo más sensible y específica posible, prefiriendo en ciertos casos que la técnica sea más sensible que específica, ya que es preferible un falso positivo a un falso negativo, ejemplo de esto es el examen para VIH en el cual es mejor tener un resultado reactivo, aunque requiera que otras técnicas más específicas confirmen o descarten este resultado, a tener un negativo falso que implique que un paciente infectado siga propagando la infección por desconocimiento.

### 2.1.2 Fundamentos de Inmunología

El sistema inmune tiene varias formas para hacer frente a los patógenos y/o cuerpos extraños que podrían significar un peligro, dentro de estas vías de defensa se encuentra la generación de anticuerpos contra antígenos específicos. Para entender esto es necesario entender ciertos conceptos:

1. Antígeno: Es cualquier sustancia que puede de generar una respuesta mediada por anticuerpos (Koivunen M., y Krogsrud R., 2006), como por ejemplo proteínas, cadenas de azúcares, etc. Dentro de esta definición se debe aclarar que existen antígenos propios, que sirven para reconocer las propias células y su condición, pero que, en condiciones normales, no generan una respuesta inmune, y los antígenos externos que sí generan una respuesta inmune como fiebre, inflamación, irritación, etc.
2. Anticuerpo: es la forma genérica de nombrar a todas las inmunoglobulinas, éstas son desarrolladas por los Linfocitos B y pueden estar adheridas a su membrana o presentes en suero y otros fluídos corporales, éstas dependiendo de su estructura espacial y por su conformación proteíca se dividen en 5 clases: IgG, IgA, IgM, IgD e IgE, en orden de mayor a menor concentración (Tabla 1). Su unidad básica tiene la foma tipo “Y” (Figura 1), donde la zona superior presenta glicoproteínas hipervariables que son las que le permiten el reconocimiento antigénico (Palomo I., y cols., 2002).

|  |
| --- |
| Tabla 1: Características de las inmunoglobulinas (Palomo I., y cols., 2002) |
|  |
| Figura 1. Esquema de la estructura básica de una Inmunoglobulina (Modificado de Koivunen M., y Krogsrud R., 2006) |
| 2017_05_01_13_51_23_Principles_of_Immunochemical_Techniques_.pdf_Foxit_Reader.png |

La capacidad de los Linfocitos B para generar anticuerpos específicos, ha permitido el desarrollo de anticuerpos monoclonales que por sí solos o de forma más compleja son las bases de todas las técnicas inmunoquímicas.

**Pruebas de laboratorio.**

Como ya se ha mencionado existen muchas técnicas que se basan en las reacciones Antígeno-Anticuerpo que difieren en su complejidad, capacidad de entregar información cualitativa, cuantitativa o semi-cuantitativa. Dentro de estas tenemos a las pruebas de VDRL e Inmunofluorescencia, entre muchas otras, que pese a su relevancia médica y cantidad de muestras procesadas, aún son procesadas de manera completamente manual.

**Inmunofluorescencia.**

La Inmunofluorescencia se basa en el uso de un anticuerpo, dirigido a al antígeno que se desea pesquisar, unido a una molécula fluorescente, llamado genéricamente fluorocromo, que frente a ciertas longitudes de onda de luz se excita y emite un brillo en otra longitud de onda particular, que es lo que se observa como resultado final, la Figura 2 grafica el comportamiento de 5 fluorocromos distintos sometidos a una luz de 488nm (luz en el espectro del color azul), mostrando que el fluorocromo FITC presenta una luminiscencia de unos 530nm (espectro de color Verde), el PE presenta una luminiscencia de 582nm (espectro de color amarillo), el PerCP-Cy5.5, una luminiscencia de 695nm (espectro de color rojo), y el PE-Cy7, una de 780nm (espectro de color infrarrojo), por su parte el fluorocromo indo-1 hi no presenta luminiscencia, ya que no es excitado por la luz a los 488nm utilizada para este ejemplo.

|  |
| --- |
| Figura 2. Gráfico comparativo de 5 fluorocromos distintos bajo una luz de 480nm. Realizado con la herramienta online de BD Biosciences. |
|  |

Existen 2 tipos de inmunofluorescencia, la inmunoflourescencia directa (IFD) y la inmunofluorescencia indirecta (IFI), en el esquema de la Figura 3 se puede apreciar la diferencia entre los 2 tipos. En el caso de IFD, el anticuerpo con el marcador flourescente va directamente unido a la célula infectada mientras que en la IFI, existe un segundo anticuerpo que es el que estará unido a la célula, y a éste se le unirá el anticuerpo unido al marcador fluorescente.

|  |
| --- |
| Figura 3. Ejemplo de IFD (I) e IFI (II). (Koivunen M., y Krogsrud R., 2006) |
|  |
| **Izquierda**: Anticuerpo Anti-A con el marcador fluorescente unido selectivamenre al Antígeno A (IFD).  **Derecha**: Anticuerpo de conejo Anti-A unido selectivamente al Antígeno A, luego un anticuerpo de de cabra con ell marcador fluorescente se unen al anticuerpo de conejo (IFI). |

Estos 2 tipos de Inmunofluorescencia se utilizan para propósitos particulares, ya que por sus características pueden entregar datos distintos. En el caso de la IFD, nos permite conocer si determinado antígeno se encuentra en las células, esto es útil para el reconocimiento de líneas celulares o si existe un determinado virus en estas células (Figura 4). Por otra parte la IFI, permite conocer si un determinado anticuerpo está presente, y además entregar información sobre en qué posición de la célula se encuentra unido los anticuerpos del pacientes, esto es patente en el caso de los anticuerpos generados en enfermedades autoinmunes donde el lugar hacia donde se dirigen los anticuerpos es indicativo de la patología presente. En la Figura 5 se observa un ejemplo de lo anterior, en el que se observan 3 patrones de positividad distintos, A) Rugoso, B) Moteado y C) Liso, lo que acompañado al criterio médico, se usará para determinar a qué tipo de enfermedad autoinmune corresponde.

|  |
| --- |
| Figura 4 Se muestra un examen de Inmunofluorescencia Directa (IFD), en el cual las células que brillan de color verde indican la presencia del Virus Respiratorio Sincicial (Vircell Microbiologists) |
| http://www.vircell.com/media/filer_public/c9/4e/c94efa9f-ce63-45d8-8167-478c1c914025/rsv_acresi.jpg |

|  |
| --- |
| Figura 5. Patrones de unión en Inmunofluorescencia Indirecta (IFI). (Koivunen M., y Krogsrud R., 2006) |
|  |

### 2.1.3. Enfermedades Autoinmunes.

Las enfermedades autoinmunes corresponden a una serie de patologías caracterizadas por la presencia de anticuerpos que atacan a distintos antígenos presentes en las propias células de la propia persona, afectando a órganos específicos, sistemas completos o diversos tejidos a lo largo de todo el cuerpo del individuo (Rosenblum, M., Remedios, K., y Abbas A. 2015), comprometiendo su salud y requiriendo largos tratamientos que la mayoría de las veces son de por vida. Ejemplo de este tipo de patologías es la Diabetes de tipo 1, en donde anticuerpos atacan los islotes de Langerhans, presentes en el páncreas, que son los responsables de la producción de insulina y el Lupus Eritematoso Sistémico, que como su nombre lo indica, afecta a diversos sistemas del cuerpo, dañando múltiples órganos.

Respecto a los exámenes realizados para pesquisar estas enfermedades, existen diversas alternativas como el Enzima Inmuno Ensayo (ELISA), sin embargo ninguno es capaz de entregar información sobre el tipo, distribución y cantidad relativa, de la presencia o ausencia de alguno de estos Autoanticuerpos que la técnica de Inmunofluorescencia Indirecta (IFI) descritas+ anteriormente. Sin embargo, el principal problema con respecto a la realización de esta IFI, radica en que su interpretación es muy subjetiva, y requiere la presencia de un personal altamente calificado para que el examen aporte toda la información que es posible extraer de su interpretación. Esta dependencia, provoca que exista falta de consistencia en los resultados informados dese uno u otro laboratorio, así como una baja eficiencia en la cantidad de muestras médicas procesadas (Gao, Z. y cols, 2015).

## 2.2. Reconocimiento de imágenes.

El reconocimiento de patrones en imágenes es una disciplina científica que busca la clasificación las imágenes en categorías o clases, cuya aplicación práctica va desde el reconocimiento de caracteres alfanuméricos hasta el de imágenes más complejas como lo son los rostros humanos. Sin embargo, uno de sus usos más interesantes está en el diagnóstico asistido por computadoras de imágenes médicas.

Desde la década de 1970, se han desarrollado técnicas de reconocimiento de imágenes de carácter médico utilizando como principio el proceso de pixeles a bajo nivel y el uso de modelos matemáticos geométricos, para construir sistemas basados en reglas que eran capaces de resolver parcialmente tareas determinadas.

La revolución en el reconocimiento de imágenes vino de la mano del desarrollo del aprendizaje de máquina (Deep Learning), que permite desarrollar modelos multicapas que permiten a los sistemas informáticos aprender a través de representaciones abstractas de la información obtenida de las imágenes procesadas, reconocer características de éstas que se utilizarán para clasificar nuevas imágenes (LeCun, Y. y colaboradores, 2015)

Los algoritmos de aprendizaje profundo (Deep Learning) mayormente utilizados usan como base las redes neuronales, que en términos simples son algoritmos de regresión lineal donde una neurona es activada recibiendo ciertos parámetros de tendencia y peso. Estas neuronas son la base de la mayoría de las metodologías de Deep Learning (Litjens, G. y cols., 2017), dentro de las cuales tenemos entre otros:

1. Redes Neuronales Convolucionales (CNN): poseen la ventaja de ser ideales para el desarrollo de algoritmos de aprendizaje sobre imágenes ya que requieren poca supervisión durante su entrenamiento y requieren el uso de múltiples neuronas con funciones particulares. Estas neuronas se agrupan en capas donde el valor de los pixeles de un sector son agregados en una función de permutación invariante, usualmente el promedio de los valores o el máximo de ellos, conservándose éste, con lo que se consigue reducir la cantidad de parámetros de la red. Estas redes terminan en una capa final que aplica una función del tipo Softmax, que genera el valor de entrenamiento de la red neuronal convolucional
2. Redes Neuronales Recurrentes (RNN): Su desarrollo fue orientado en una primera instancia para el análisis de secuencias discretas. La conexión entre las unidades que la conforman forma un ciclo dirigido, lo que permite que exhiban un comportamiento temporalmente dinámico. Su uso actual se encuentra en el reconocimiento de escritura manuscrita y en la traducción de textos.
3. Codificadores automáticos (AEs) y Codificadores automáticos apilados (SAEs): estas se conforman de redes simples, y están entrenadas para reconstruir la entrada X en la capa de salida x’, a través de una capa oculta intermedia h. Habitualmente para su uso predictivo son entrenadas de manera individual para luego ajustar la red.
4. Maquinas Boltzman restringidas (RBMs): está constituido de una capa visible y otra oculta y sus nodos son bidireccionales, por lo cual pueden obtener a partir de una entrada de vector X, una salida H y viceversa.

Litjens, G. y cols. (2017) desarrolló una revisión de diversos papers desarrollados desde 2012 sobre el uso de distintas técnicas de Deep Learning en el reconocimiento de imágenes de origen médico. En la Figura 6 podemos observar que el peak de papers de investigación respecto a este tema es este año, pudiéndose ver que sólo considerando las investigaciones publicadas hasta junio de este año, ya se han realizado más investigaciones que durante todo el 2016. Otro punto relevante de esta figura es que la técnica de Deep Learning usada preferentemente por todos los investigadores fueron las Redes Neuronales Convolucionales (CNN), en lugar de otras técnicas existentes.

|  |
| --- |
| Figura 6. Número de papers sobre Deep Learning por técnica utilizada, realizados entre 2012 y Junio de 2017. (Litjens, G. y cols., 2017) |
|  |

## 2.3. Reconocimiento del problema, necesidad y oportunidad.

**Los 5 por qué.**

Se desarrolla, a través de la técnica de “los 5 por qué”, el reconocimiento del problema u oportunidad:

1: ¿Por qué la IFI para Anticuerpos Autoinmunes es una técnica compleja de informar?

Porque los patrones de reactividad son complejos de interpretar.

2: ¿Por qué los patrones de reactividad son complejos de interpretar?

Porque su correcta interpretación depende de la experiencia de quien informa.

3: ¿Por qué su correcta interpretación depende de la experiencia de quien informa?

Porque es frecuente que existan discrepancias en la interpretación de los exámenes entre distintos profesionales.

4: ¿Por qué es frecuente que existan discrepancias en la interpretación de los exámenes entre distintos profesionales?

Porque cada profesional tiende a informar basado en su propia subjetividad.

5: ¿Por qué cada profesional tiende a informar basado en su propia subjetividad?

Porque no existe un método implementado que permita objetivar los resultados de las IFI.

|  |
| --- |
| Figura 7. Ishikawa del problema. Elaboración propia |
|  |

Tomando en consideración “Los 5 Por qué” y el problema considerado en el Ishikawa de la Figura 7 se debe desarrollar una solución que permita al personal contar con una segunda opinión objetiva adicional a su propia opinión, ayudando así a los profesionales en su labor diaria, en especial a aquellos con menos experiencia, disminuyendo la probabilidad de error del personal. El funcionamiento de esta herramienta de apoyo debe considerar que todo resultado debe pasar por el criterio del personal antes de ser informado. Esta herramienta se hará cargo de los puntos 2, 3 y 5 del Ishikawa, como se trata de un sistema de segunda opinión deberá reconocer los patrones de Inmunofluorescencia (punto 2), además deberá alcanzar un grado de confiabilidad adecuado para ser un apoyo al personal con menos entrenamiento (punto 3), además debe respaldar los resultados interpretados para ser consultados en caso de necesidad de una nueva observación (punto 5).

### 2.3.1 Posibles soluciones

En la Tabla 2 se observan distintas soluciones a los problemas y necesidades detectadas en el punto anterior.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tabla 2. Comparación de diversas soluciones (Elaboración propia) | | | | | |
|  | **Beneficio para Profesional** | **Requiere mayor experiencia** | **Requiere coste mayor** | **Permite reducir la probabilidad de error** | **Observaciones** |
| **Uso de Mejor Microscopio** | Mejora la visualización de los resultados | No | Sí (++) | Probablemente | Dependiendo de las condiciones podría no tener un gran impacto. |
| **Capacitación más extensa** | Mejora la interpretación por parte del profesional | Sí | Sí (++) | Sí | Aumenta el costo para el laboratorio.  Difícilmente se podría conseguir el mismo nivel para todos los profesionales  Requiere mucho tiempo para alcanzar este objetivo |
| **Uso de ELISA o WesternBlot** | Técnica Semi-Automatizada | Poca | Sí (+++) | Depende | Estás técnicas suelen ser muy caras y aunque pueden ser específicas para algunos anticuerpos, no entregan la misma información que la IFI |
| **Equipo de interpretación** | Mejora la probabilidad de reconocer los patrones | No | Sí (+) | Sí, probablemente | Un equipo de trabajo ayudaría a interpretar mejor los resultados, pero implicaría un mayor staff trabajando sólo en interpretación lo cual es más caro e impide una mejor distribución del personal |
| **Sistema que reconozca los patrones** | Mejora la interpretación | Poca | Si(+) | Sí | Es una opinión adicional al criterio profesional  Ayuda al personal con menos experiencia. |

**2.3.2. Alternativas de solución ideal.**

La solución ideal para este proyecto se debe basar en la automatización de parte o total del proceso de las técnicas en cuestión, sin embargo el procesamiento de las muestras requeriría de la elaboración de maquinarias especializadas de alto coste, por lo que este proyecto se centrará en la búsqueda de alguna alternativa que permita la automatización de la lectura e interpretación de estos exámenes. Con este objetivo se deberá evaluar las alternativas de código inteligente disponibles. La Tabla 3 es una comparativa de las técnicas que se disponen.

|  |  |
| --- | --- |
| Tabla 3. Comparativa de algoritmos inteligentes (Creación Propia) | |
| Alternativa | Observaciones |
| Árbol de Decisión | **Aprendizaje inductivo: se crea a través de propiedades dadas una respuesta inducida a cada iteración. Requiere el ingreso de patrones lo más completos posibles (**Barrientos, R. y cols., 2009) |
| Naïve Bayes | Mediante el uso de Curvas Gaussianas permite la predicción de tendencias, asociaciones y predicciones. Requiere estadística previa (Lowd, D. y Domingos, P., 2005) |
| Redes Neuronales Recurrentes (RNN) | Como ya se mencionó antes su principal uso se encuentra en el reconocimiento de escritura manuscrita y en la traducción de textos. |
| Redes Neuronales Convolucionales (CNN) | Mediante el uso de Redes Convolucionales, permite que a partir de ejemplos dados, el sistema aprenda a responder a ellos, generando por sí mismo los patrones requeridos. (Schmidhuber, J., 2008.) |

Cómo se puede observar de la Tabla 3 las 2 primeras alternativas son ejemplos de aprendizaje y predicción de patrones, sin embargo se requiere que exista un estudio previo y se utilicen una gran cantidad de recursos previos antes de su implementación y aún así sería difícil obtener resultados que pudieran satisfacer la necesidad de reemplazar la experticia humana por lo que de aquí inducimos que por ser una experiencia nueva de la cual no existen ni se basan en datos estadísticos, o que permitan establecer patrones previos, sería útil el uso de alguna técnica de Deep Learning, por ello se decide el uso de Redes Neuronales Convolucionales, que además de tener una utilidad en la generación de modelos de reconocimiento de imágenes, ya se ha establecido como la técnica de elección en las imágenes de origen médico.

### 2.3.3 Solución Propuesta

Basados en los antecedentes previamente detallados, se propone la implementación de un sistema que mediante el uso de Redes Neuronales Convolucionales sea capaz de clasificar las muestras obtenidas ópticamente con una precisión buena precisión y que funcione como apoyo en la decisión del informe que generará el Profesional de Laboratorio a cargo de la realización de las Inmunofluorescencia para Anticuerpos Autoinmunes.

Debido a la necesidad de procesar las imágenes de las muestras obtenidas a través de un entorno Linux, y ya que el gran general de los sistemas de informe de exámenes están realizados en entorno Windows, se propone que el sistema funcione en parte desde una caja de procesamiento en un entorno Linux, y que esta caja sea el intermediario entre el modelo de reconocimiento y la interfaz de operación del usuario, en la Figura 8 se observa un esquema de como trabajaría el sistema.

|  |
| --- |
| Figura 8. Propuesta del sistema (Elaboración propia) |
|  |

## 2.4. Objetivos

**Objetivo General**

Generar un sistema que apoye a los Profesionales de laboratorio en la clasificación de las pruebas para Inmunofluorescencia para Auto-anticuerpos.

**Objetivos específicos**

* Generar un modelo de Redes Neuronales Convolucionales que reconozca los resultados Negativos y los Patrones Homogéneo y Moteado.
* Determinar en menos de 1 minuto si una muestra es positiva o negativa
* Clasificar automáticamente las muestras positivas en Patrón Homogéneo o Patrón moteado.
* Presentar tanto imágenes como resultados de manera que el profesional a cargo pueda decidir si modifica o no la clasificación realizada por el sistema.
* Cargar nuevos pacientes y muestras.

## 2.5.- Alcance:

1. Reconocer patrones de propios de las pruebas de Inmunofluorescencia para Anticuerpos Autoinmunes.
2. Reconocer Resultados Negativos y Patrones Homogéneo y Moteado.
3. Reconocer imágenes provenientes de muestras de provenientes de medios externos al sistema.

## 2.6. Limitaciones

* No automatiza el proceso previo a la lectura
* Depende de que las imágenes posean una relativamente buena calidad para su clasificación.
* El sistema sólo clasificará las muestras que sean Negativas o que posean los Patrones Homogéneo o Moteado, el resto de los patrones no está considerado en esta versión.
* Requiere una conexión a Internet para su correcto funcionamiento.
* Requiere una red para conectarse con el dispositivo Raspberry pi.
* El informe final y definitivo de los exámenes es responsabilidad del Profesional del Laboratorio, por lo que se requiere que el personal revise y evalúe personalmente la información entregada por el sistema informático.

**2.7. Requerimientos**

**Requerimientos de negocio:**

* RN1: Leer las pruebas de Inmunofluorescencia para Anticuerpos Autoinmunes.
* RN2: Identificar correctamente las pruebas de cada paciente para su correcto informe.
* RN3: Enviar los resultados finales para su informe.
* RN4: Guardar registro de los exámenes

**Requerimientos de usuario:**

* RU1: Utilizar las imágenes entregadas por el usuario para su clasificación.
* RU2: Reconocer los patrones de positividad/negatividad de los exámenes
* RU3: Permitir observar los resultados e imágenes para su evaluación visual.
* RU4: Presentar los resultados identificando a cada paciente.
* RU5: Enviar los resultados ya evaluados visualmente para su informe final.
* RU6: Guardar las imágenes y los resultados del equipo y evaluación para su respaldo.

**Requerimientos de Software:**

* RS1: Recibir los datos e imágenes entregadas por los usuarios
* RS2: Asignar las imágenes en una carpeta para cada petición.
* RS3: Realizar el pre-procesamiento necesario a las imágenes subidas
* RS4: Usar los modelos desarrollados para clasificar los resultados de los pacientes.
* RS5: Utilizar una interfaz capaz de mostrar los datos procesados y las imágenes por cada paciente
* RS6: Reconocer cada muestra dependiente del paciente.
* RS7: Poder diferenciar cada petición de cada orden de examen.
* RS8: Enviar los datos del examen para el informe de resultados
* RS9: Mantener un respaldo de los resultados e imágenes clasificadas.

# Capitulo III: Materiales y Métodos

**3.1 Selección de metodologías**

Para la elaboración del proyecto se opta por el uso de una metodología ágil ya que el proyecto es de carácter innovativo de tecnología aplicada a los exámenes en cuestión, por lo que se utilizará SCRUM, ya que permite regulares iteraciones que mejorarán el resultado final del software. Por su parte el desarrollo del producto se usará un desarrollo Iterativo Incremental que permitirá tanto evaluar los resultados obtenidos en cada iteración como el ajuste de estos y la implementación de características adicionales que deba tener el producto final.

Los roles asignados para la metodología SCRUM, serán:

- Product Owner: Romina Torres, Profesores guía

- Scrum Master: Romina Torres, Profesora guía

- Scrum Team: César Cheuque Cerda, Alumno Tesista

Siguiendo la metodología Scrum se procede a la realización de las Historias de Usuario, realizadas en la plantilla destinada para este fin. Las Historias de Usuario se definieron en las primeras reuniones de con el Product Owner, y se priorizaron al inicio de cada Sprint. Cada Sprint debería tomar 2 a 3 semanas, y se considerará aceptado una vez se haya completado las tareas de la Pila del Sprint.

## 3.2 Selección de técnicas y tecnologías.

Para el desarrollo de los modelos se utilizó Tensorflow como backend y Keras como API de configuración de estos modelos, todo fue realizado sobre Ubuntu 16.04 y escrito en Python 2.7. Todo esto fue implementado sobre una Laptop AMD8, sin una GPU compatible con CUDA.

El evaluador de imágenes que utiliza los modelos ya entrenados, fue desarrollado sobre una Raspberry pi 3 Modelo B, que fue la misma sobre la que se desarrolló una API Rest para controlar la Aplicación Web de la Interfaz de Usuario, para ello fue utilizado Django y las librerías de Django Rest Framework. Para la interfaz fue utilizado Bootstrap, además de las librerías de desarrollo Web DataTables y FancyBox.

**3.3 Matriz de trazabilidad de requerimientos**

En la Tabla 4 de realiza una matríz de trazabilidad de las 3 dimensiones de requerimientos del proyecto.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tabla 4.Matriz de trazabilidad de requerimientos | | | | | | |
| Requerimientos de Software | Requerimientos de Usuarios | | | | | |
| RU1 | RU2 | RU3 | RU4 | RU5 | RU6 |
| RS1 | RN1 |  |  |  |  |  |
| RS2 |  |  | RN2 |  |  |  |
| RS3 |  | RN1 |  |  |  |  |
| RS4 |  | RN1 |  |  |  |  |
| RS5 |  |  | RN2 |  |  |  |
| RS6 |  |  | RN2 |  |  |  |
| RS7 |  |  |  | RN2 |  |  |
| RS8 |  |  |  |  | RN3 |  |
| RS9 |  |  |  |  |  | RN4 |

**3.4 Planificación del proyecto**

Se realizará una metodología SCRUM, por lo que se precederá al procesamiento de varios Sprint consecutivo en los que se irá liberando código y aumentando las funcionalidades del proyecto.

El proyecto se dividió en 2 fases, la primera en la que abarcó los meses de marzo a julio de 2017, y una segunda que abarcó desde agosto a diciembre de 2017. En la primera fase se realizó un modelo de reconocimiento de imágenes de la prueba de VDRL, que es una prueba para el diagnóstico de Sífilis, pero que fue descartado a favor de un modelo para el reconocimiento de las muestras de Inmunofluorescencia para Anticuerpo Autoinmune, que fue desarrollado en la segunda fase del proyecto. A su vez, la fase 2 fue dividida en 2 partes, la primera fue el Núcleo de Reconocimiento de Imágenes que implicó el desarrollo del modelo de reconocimiento, el entrenamiento y validación, y un programa que utilice este modelo para el reconocimiento de nuevas imágenes, y la segunda parte fue el desarrollo de la interface de una Aplicación Web para la interacción con el usuario, que incluye el uso de una API Rest.

## 3.5 Métricas de los objetivos.

**Métricas de la Fase 1:**

**Líneas de Código:**

Hito 1:

Entrenamiento y validación: 0.110 KLOCS

Hito 2:

Entrenamiento y validación primer Dataset: 0.122 KLOCS

Entrenamiento y validación segundo Dataset: 0.170 KLOCS

Núcleo de Evaluación de imágenes: 0.061 KLOCS.

Hito 3:

Entrenamiento y validación primer Dataset: 0.157 KLOCS

Entrenamiento y validación segundo Dataset: 0.170 KLOCS

Núcleo de Evaluación de imágenes: 0.066 KLOCS

Núcleo de Evaluación de imágenes segundo Dataset: 0.061 KLOCS

**Número de imágenes:**

Hito 1:

Entrenamiento y validación: 9423 imágenes.

- 4711 imágenes Reactivas

- 4712 imágenes No Reactivas.

Evaluación de imágenes: 20 imágenes

Hito 2:

Entrenamiento: 11799 imágenes

- 6119 imágenes Reactivas

- 5680 imágenes No Reactivas

Validación: 11286 imágenes

- 5538 imágenes Reactivas

- 5748 imágenes No Reactivas

Evaluación de imágenes: 3995

**Métricas Fase 2:**

**Imágenes:**

* Entrenamiento y Validación:
  + Negativo: 15.000
  + Patrón Homogéneo: 17.000
  + Patrón Moteado: 19.000
* Evaluación de los modelos:
  + Negativo: 100
  + Patrón Homogéneo:100
  + Patrón Moteado: 100

**Líneas de Código:**

Entrenamiento y Validación: 0.611 KLOCS

Entrenamiento y Validación (KLOCS reutilizados): 0.140 KLOCS

Evaluación de nuevas imágenes: 0.160 KLOCS

Evaluación de nuevas imágenes (KLOCS reutilizados): 0.053 KLOCS

Desarrollo API Rest – Plataforma Web: 0.576 KLOCS

Desarrollo Templates: 0.744 KLOCS.

## 3.6 Plan de Gestión de la Configuración

Para evitar pérdidas de información y recursos desarrollados, se realiza respaldo de todo en un repositorio privado en los servidores de GitHub.

Desarrollo de modelos, entrenamiento y validación:

<https://github.com/Arkantos2040/DeepLab-INMUNO>

Desarrollo API Rest, Aplicación Web, Evaluación de imágenes:

<https://github.com/Arkantos2040/webtesis>

## 3.7 Plan Gestión de Cambios.

Para la gestión de los cambios se puede realizar a través de la generación de Tickets de cambio directamente en la página de GitHub, sin necesidad de desarrollar en este momento otro sistema para su gestión.

Al 25-12-2017 existen 3 tickets abiertos para su modificación en la página web:

<https://github.com/Arkantos2040/webtesis/issues>

**3.8 Plan de Gestión de Riesgos**

**Identificación del Riesgo Técnico.**

En la Tabla 5 se registran los riesgos técnicos identificados hasta ahora, que serán evaluados para buscar medidas de mitigación.

|  |  |  |
| --- | --- | --- |
| Tabla 5. Identificación descripción del riesgo técnico. | | |
| Identificación del riesgo | Riesgo | Descripción |
| RT1 | Pérdida de Información crítica o parte del código | Existe la posibilidad de sobreescribir código funcional y perder funcionalidad, o que por otros motivos se pierda parte fundamental de la información |
| RT2 | Falla del hardware donde operan los sistemas. | Existe la posibilidad de que fallas en los componentes o el software sobre el que operan los programas presente fallos |
| RT3 | Falta de experiencia en el desarrollo de alguno de los sistemas necesarios para el sistema. | Existen tecnologías sobre las cuales no se tiene dominio o conocimiento y podrían implicar retrasos o problemas en algún punto del desarrollo |
| RT4 | No contar con las imágenes necesarias para el desarrollo del proyecto | Durante el desarrollo del proyecto será necesario el uso de imágenes para el entrenamiento neuronal, o para la validación de la técnica como funcional. |
| RT5 | El uso de imágenes podría implicar aspectos legales que fueran necesarios de cubrir. | El uso de las imágenes pudiera necesitar cumplir ciertos requisitos para su uso o para su protección. |

**Planificación y mitigación de los riesgos técnicos**

En la Tabla 6 se realiza la planificación de los riesgos descritos y se establecen los planes de mitigación correspondientes.

|  |  |  |
| --- | --- | --- |
| Tabla 6. Planificación de medidas de mitigación del riesgo técnico | | |
| Identificación del riesgo | Planificación | Medidas de mitigación |
| RT1 | Investigar las plataformas para el respaldo de información y el control de versión de código | Respaldo de los códigos y documentación desarrollados en un repositorio privado en GitHUB |
| RT2 | Investigar medidas de derivación para el desarrollo correcto del proceso | Se desarrolla el código orientado al uso de CPU, por lo que no requiere hardware complejo, además de desarrollar documentación adecuada para el proceso de instalación del software requerido. Se comprueba la factibilidad de correr el desarrollo en un servidor AWS |
| RT3 | Se realiza investigación de las tecnologías necesarias y su documentación | Se solicita capacitación a personas con conocimientos previos de estas tecnologías. Se prefiere el uso de tecnologías con buena cantidad de documentación |
| RT4 | Se realiza búsqueda de Dataset públicos y privados de estas imágenes y de laboratorios que realicen estas pruebas | Se logra compromiso de ayuda de 2 laboratorios privados y un laboratorio de un Hospital público. |
| RT5 | Buscar información legal, evitar que las imágenes sean enviadas a Internet. | Las imágenes no serán asociadas a ningún paciente ni serán publicadas por lo que no debiera haber problemas legales |

**3.9 Ambiente de desarrollo, pruebas y producción**

**Ambiente de Desarrollo**

El ambiente de desarrollo será Ubuntu 16.04 corriendo en un computador sin GPU nVidia.

**Ambiente de Pruebas.**

En una primera instancia las pruebas se desarrollaran en un ambiente Ubuntu, luego de lo cual se desarrolla una plataforma de trabajo en una Raspberry pi 3 Modelo b. Esta última será modificada en función de las necesidades de procesamiento de las pruebas realizadas.

**Ambiente de Producción**

Se prevee que se seguirá utilizando la Raspberry pi 3 Modelo b, durante la producción del producto, sin embargo también dependiendo la potencia necesaria se optará por otros equipos o plataformas.

# Capitulo IV: Desarrollo y Resultados

## 4.1 Desarrollo de los Sprint.

Como ya se definió anteriormente se desarrollaron las Historias de Usuario que se requieren desarrollar en este proyecto, en la Tabla 7 se observa las Historias de Usuario junto a la priorización por la técnica MOSCOW.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tabla** 7**. Historias de Usuario** | | | | | |
| **Backlog ID** |  | **Usuario** | **Descripción** | **Priozación** | **Estado** |
| **1** | yo como | **tecnólogo médico** | Quiero obtener una segunda opinión mientras realizo la inmunofluoresencia para reducir la probabilidad de error humano | S | Realizado |
| **2** | yo como | **tecnólogo médico** | Quiero que los exámenes sean fáciles de evaluar | S | Realizado |
| **3** | yo como | **tecnólogo médico** | Quiero tener seguridad de que la segunda opinión es confiable para utilizarla como apoyo a mis decisiones | M | Realizado |
| **4** | yo como | **tecnólogo médico** | Quiero validar que es posible clasificar automáticamente de las imágenes de imunofluorescencia | M | Realizado |
| **5** | yo como | **tecnólogo médico** | Quiero trabajar con un servicio web que me permita utilizar el sistema apropiadamente | M | Realizado |
| **6** | yo como | **tecnólogo médico** | Quiero trabajar con una interface operativa | S | Realizado |
| **7** | yo como | **laboratorio** | Quiero virtualizar el servicio para poder venderlo para otros laboratorios | W | Pendiente |
| **8** | yo como | **laboratorio** | Quiero asegurarme de que otros no puedan acceder a los datos confidenciales para evitar demandas | M | Realizado |
| **9** | yo como | **laboratorio** | Quiero asegurarme de que los resultados se podrán traspasar automáticamente al sistema de informes | C | Pendiente | |

A continuación se despliegan los Sprint realizados.

**Sprint 1**: Historia de Usuario 5.

**Nombre del Sprint:** Quiero trabajar con un servicio web que me permita utilizar el sistema apropiadamente

**Objetivo sprint**: Desarrollo de un Servicio Web que mediante el uso de una API Rest me permita controlar las conexiones con los distintos componentes.

Para desarrollar el Sprint se define una serie de tareas que forman parte de la Pila del Sprint de esta Historia de Usuario (Tabla. 8).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tabla 8. Pila del Sprint 1, la que sigue los pasos para realizar la Historia de Usuario 5. | | | | | |
| **Backlog ID** | **Tarea** | **Tipo** | **Estado** | **Fechas** | **Estimación** |
| **1** | Creación de Casos de Uso | Modelado | Realizado | César Cheuque C. | 1 |
| **2** | Creación del Diagrama de Despliegue | Modelado | Realizado | César Cheuque C. | 1 |
| **3** | Modelado de la Base de Datos | Modelado | Realizado | César Cheuque C. | 3 |
| **4** | Desarrollo del Servidor con Apache | Codificación | Realizado | César Cheuque C. | 3 |
| **5** | Creación de Servidor Web con Django | Codificación | Realizado | César Cheuque C. | 10 |
| **6** | Conexión a Base de Datos | Codificación | Realizado | César Cheuque C. | 5 |
| **7** | Creación de la API REST | Codificación | Realizado | César Cheuque C. | 10 |
| **8** | Comunicación con las Aplicaciones Python | Codificación | Realizado | César Cheuque C. | 2 |
| **9** | Pruebas de Comunicación entre componentes | Codificación | Realizado | César Cheuque C. | 10 |

Tareas del Sprint.[[1]](#footnote-1)

**1. Creación de Casos de Uso: En la Figura 9 se despliega el Diagrama de casos de uso del Usuario del Sistema:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Figura** 9**.** Diagrama de Casos de Uso | | | |
|  | | | |
| CU: 1 | | | |
| Nombre: | Ingreso al sistema | | |
| Descripción | El usuario hace login con su nombre y password | | |
| Actor: | Usuario | | |
| Acción: | ACTOR | SISTEMA |  |
|  | 1. Ingresa USUARIO 2. Ingresa Clave | 1. Valida existencia del usuario. 2. Valida que el password corresponde al usuario 3. Accede a la pantalla de resultados |  |
| CU: 2 | | | |
| Nombre: | Subir de exámenes | | |
| Descripción | Se ingresa nueva petición para el procesamiento de las nuevas muestras | | |
| Actor: | Usuario | | |
| Acción: | ACTOR | SISTEMA |  |
|  | 1. Presiona botón de Ingreso de Nuevas Muestras | 1. Valida que se ingresen todos los parámetros. 2. Llama al programa que inicia el procesamiento de exámenes 3. Ordena las imágenes y actualiza la base de datos. |  |
| CU: 3 | | | |
| Nombre: | Revisión de los resultados | | |
| Descripción | El usuario revisa los resultados entregados por el sistema | | |
| Actor: | Usuario | | |
| Acción: | ACTOR | SISTEMA |  |
|  | 1. Selecciona una determinación en el menú de muestras. | 1. Despliega las imágenes procesadas 2. Despliega el resultado de la clasificación |  |
| CU: 4 | | | |
| Nombre: | Corrección del número de Orden Médica | | |
| Descripción | El usuario modifica el número de orden asociado a la determinación. | | |
| Actor: | Usuario | | |
| Acción: | ACTOR | SISTEMA |  |
|  | 1. Selecciona una determinación en el menú de muestras. 2. Presiona el botón de Modificación de muestra.   4. Ingreso del nuevo número  6. Confirma acción | 1. Despliega casilla de ingreso de datos   5. Solicita confirmación  de la acción  7. Actualiza la base de  datos |  |

|  |  |  |  |
| --- | --- | --- | --- |
| CU: 5 | | | |
| Nombre: | Corrección de Resultados | | |
| Descripción | El usuario modifica, basado en las imágenes desplegadas, el resultado de la clasificación automática. | | |
| Actor: | Usuario | | |
| Acción: | ACTOR | SISTEMA |  |
|  | 1. Selecciona una determinación en el menú de muestras. 2. Presiona el botón de Modificación de Resultados.   4. Ingreso nueva  clasificación  6. Confirma acción | 1. Despliega casilla de selección de resultado   5. Solicita confirmación  de la acción  7. Actualiza la base de datos |  |

|  |  |  |  |
| --- | --- | --- | --- |
| CU: 6 | | | |
| Nombre: | Envío de exámenes para informes | | |
| Descripción | El usuario envía los resultados de la clasificación para su informe | | |
| Actor: | Usuario | | |
| Acción: | ACTOR | SISTEMA |  |
|  | 1. Selecciona una determinación en el menú de muestras. 2. Presiona el botón de Envío de resultados | 1. Valida existencia del número de Orden Médica. 2. Envía los resultados al LIS 3. Actualiza datos de la base de datos. |  |

|  |  |  |  |
| --- | --- | --- | --- |
| CU: 7 | | | |
| Nombre: | Respaldo de exámenes | | |
| Descripción | Se saca del sistema una muestra para que no pueda ser editada | | |
| Actor: | Usuario | | |
| Acción: | ACTOR | SISTEMA |  |
|  | 1. Selecciona una determinación en el menú de muestras. 2. Presiona el botón de Respaldo   4. Confirma Respaldo | 1. Solicita confirmación de la acción   5. Actualiza la base de  datos.  6. Actualiza pantalla de  resultados |  |

|  |  |  |  |
| --- | --- | --- | --- |
| CU:8 | | | |
| Nombre: | Cerrar sesión | | |
| Descripción | Se cierra sesión y se regresa al menú de login | | |
| Actor: | Usuario | | |
| Acción: | ACTOR | SISTEMA |  |
|  | 1. Presiona el botón de Cierre de Sesión   3. Confirma acción. | 1. Solicita confirmación de la acción   4. Cierra sesión  5. Abre login del Sistema |  |

**2. Creación del Diagrama de Despliegue**

En la Figura 10, se presenta el diagrama de despliegue del sistema, en la cual se observa que se dividirá en la Estación de trabajo, que es por donde interactuará el usuario con el Sistema, el Servidor de Base de Datos que contiene la base de datos MySQL del Sistema, el Core de Aplicaciones Python que serán las encargadas de interactuar con las muestras y su clasificación. Finalmente tenemos al Web Server que mediante REST, comunica e interactúa con todos los componentes del Sistema.

|  |
| --- |
| Figura 10. Diagrama de Despliegue del Sistema |
|  |

**3. Modelado de la Base de Datos**

En la Figura 11 se presenta el modelo de la base de datos del sistema. Django Rest Framework crea dentro de sí el diseño del modelo de la base de datos, generando para sí mismo el usuario y los permisos para cada uno, por lo que en realidad esta tabla no se creará, pero sí se dejará explicitado.

|  |
| --- |
| Figura 11. Modelo de Base de Datos. Elaboración propia |
|  |

**4. Desarrollo del Servidor**

Para el desarrollo del Servidor Web, se optó por desarrollarlo sobre la Raspberry pi 3B, por lo que se procede a instalar lo necesario para el correcto funcionamiento del sistema:

- Apache 2:

$ sudo apt-get install apache2

- Librerías para soporte de PHP:

$ sudo apt-get install php5 libapache2-mod-php5

- MySQL:

$ sudo apt-get install php5-mysql mysql-server mysql-client

- phpMyAdmin:

$ sudo apt-get install phpmyadmin

Finalmente se agrega la línea “Include /etc/phpmyadmin/apache.conf” al final del archivo /etc/apache2/apache2.conf y se reinicia Apache.

**5. Creación de Servidor Web con Django**

Se opta por usar Django como base para el Servicio Web por ser dúctil, seguro, además de que se puede usar con Python sin necesidad de sobrecargar demasiado a la Raspberry, además que posee una buena documentación y una comunidad que lo utiliza habitualmente

**6. Conexión a Base de Datos**

Para esto se procede a instalar un conector para poder utilizar la base de datos creada en MySQL con el entorno Python.

$ Sudo pip install mysql-connector==2.1.4

**7. Creación de la API REST**

Al estar trabajando con Django, se opta por usar una herramienta que facilita la implementación de este tipo de API, este se llama Django REST Framework y para su instalación se siguieron los siguientes pasos:

$ pip install djangorestframework

$ pip install markdown

$ pip install django-filter

**8. Comunicación con las Aplicaciones Python**

Se terminará en la fase de integración, pero en pruebas preliminares se ha logrado conseguir el funcionamiento de scripts Python desde la web.

**9. Pruebas de Comunicación entre componentes**

Estas pruebas se realizan directamente al momento de realizar la página web

**Criterios de Aceptación del Sprint:**

- Se logra conexión con el servidor desde otro computador de la red.

- Se logra conexión con la base datos, se puede editar su contenido.

Problemas y desafíos:

- Respaldar la información y base de datos.

- Desarrollar componentes que puedan conectarse con un Sistema de informes de laboratorio (LIS).

Durante el desarrollo del Sprint 1, hubo una serie de complicaciones que fueron desde problemas para configurar los parámetros, instalar componentes hasta tener que partir de cero debido a errores con la tarjeta de memoria. Esto llevó a más de duplicar el tiempo estimado para este Sprint. En la Figura 12 se presenta el Burndown Chart de este Sprint.

|  |
| --- |
| Figura 12. Burndown Chart del Sprint 1. |
|  |

**Sprint 2: Historia de Usuario 6.**

**Nombre del Sprint**: Quiero trabajar con una interface operativa

**Objetivo sprint:** Desarrollo de una Interfaz Web que permita al usuario del sistema utilizar las funciones previamente desarrolladas en la API Web hecha con Django Rest Framework.

La pila del Sprint de desarrollo de la Historia de Usuario 6, se despliega en la Tabla 9.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tabla 9. Pila del Sprint 2, la que sigue los pasos para realizar la Historia de Usuario 6. | | | | | |
| **Backlog ID** | **Tarea** | **Tipo** | **Estado** | **Fechas** | **Estimación** |
| **1** | Preparación de los Mockups | Modelado | Realizado | César Cheuque C. | 5 |
| **2** | Buscar Alternativas de desarrollo | Investigación | Realizado | César Cheuque C. | 10 |
| **3** | Aprendizaje de HTML, CSS, y JS | Investigación | Realizado | César Cheuque C. | 30 |
| **4** | Creación de las Web basadas en Mockups y la API Desarrollada | Codificación | Realizado | César Cheuque C. | 30 |
| **5** | Ajustes de la API | Codificación | En Proceso | César Cheuque C. | 10 |

Desarrollo del Sprint 2:

**1. Preparación de los Mockups:**

Se desarrolla los Muckups de todas las pantallas que contendrá la Web con ayuda de Powerpoint para brindarle cierta capacidad de respuesta frente a la operación de algunos testers. Se mostró esta “DEMO” a 4 Tecnólogos Médicos para ver su respuesta a la Web y tomar nota de las sugerencias que estos profesionales realizan.

**2. Buscar Alternativas de desarrollo Web.**

Debido a las características propias de la plataforma en que se realizará esta Aplicación Web, se buscó una alternativa que fuera liviana y que fuese compatible con Django. Es por esto que se escogió el uso de Bootstrap, que es compatible y posee una buena cantidad de documentación.

**3. Aprendizaje de HTML, CSS, y JS.**

Durante el desarrollo de este Sprint fue necesario aprender sobre el funcionamiento de una Web, logrando realizar una aplicación Web funcional con las capacidades y funciones esperadas.

**4. Creación de las Web basadas en Mockups y la API Desarrollada**

La realización de la Aplicación Web, fue desarrollada directamente sobre la Api creada en Django como Template de este. Lo que ayuda a que pueda consumir los servicios de esta API. En la Figura 13, se muestra la estructura final del proyecto. Las 2 páginas principales de la Aplicación Web, se muestran en las Figuras 14 y 15.

|  |
| --- |
| Figura 13. Estructura del Proyecto. |
|  |
| Figura 14. Página de Login del Sistema |
|  |

|  |
| --- |
| Figura 15. Página principal de la Aplicación Web. |
|  |

**5. Ajustes de la API.**

El Ajuste final de la API Rest, es un proceso que depende de las necesidades que han aparecido durante la implementación de la Web y por tanto aunque ahora esté funcionando la API Rest, es posible que existan nuevas modificaciones si así se requiera.

**Criterios de Aceptación del Sprint:**

- Se crea la web usando directamente la API desarrollada desde la Web.

- La Aplicación Web, logra funcionar de acuerdo a sus funciones requeridas

- La web app es capaz de desplegar la información e imágenes de las peticiones

**Problemas y desafíos:**

- Mostrar las Imágenes correctamente.

- Crear un respaldo adecuado para las Imágenes y resultados

- Integrar nuevas funciones

Durante el desarrollo del Sprint 2, al ser la primera vez que se realizaba un proyecto de este tipo, se requirió un tiempo mayor del que se estimaba, por lo que eso, sumado a poder trabajar adecuadamente con la API implementada anteriormente, llevaron a seguir trabajando en esta web, aún pasado el tiempo destinado para el Sprint. En la Figura 16 se presenta el Burndown Chart de este Sprint.

|  |
| --- |
| Figura 16. Burndown Chart del Sprint 2. |
|  |

**Sprint 3: Historia de Usuario 4.**

**Nombre del Sprint**: Quiero validar que es posible clasificar automáticamente de las imágenes de imunofluorescencia

**Objetivo sprint:** Desarrollar un modelo de reconocimiento de imágenes de Inmunofluorescencia que pueda clasificar las muestras en Negativo, Patrón Homogéneo o Patrón Moteado.

La pila del Sprint de desarrollo de la Historia de Usuario 4, se despliega en la Tabla 10.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tabla 10. Pila del Sprint 3, la que sigue los pasos para realizar la Historia de Usuario 4. | | | | | |
| **Backlog ID** | **Tarea** | **Tipo** | **Estado** | **Fechas** | **Estimación** |
| **1** | Obtención de Imágenes | Trabajo de Campo | Realizado | César Cheuque C. | 10 |
| **2** | Clasificación visual de las imágenes | Clasificación | Realizado | César Cheuque C. | 40 |
| **3** | Aprendizaje de modelos de clasificaciones múltiples | Investigación | Realizado | César Cheuque C. | 7 |
| **4** | Creación de algoritmos de Clasificación | Codificación | Realizado | César Cheuque C. | 10 |
| **5** | Entrenamiento de las Redes Neuronales. | Iteración | Realizado | César Cheuque C. | 40 |
| **6** | Ajustes de modelos y algoritmos | Codificación | Realizado | César Cheuque C. | 20 |
| **7** | Realización de Algoritmo de Evaluación de Imágenes | Codificación | Realizado | César Cheuque C. | 5 |
| **8** | Pruebas de Clasificación | Pruebas | Realizado | César Cheuque C. | 3 |

**Desarrollo del Sprint 3.**

**1. Obtención de Imágenes**

La obtención de imágenes fue un problema debido a que no todos los laboratorios que realizan la técnica poseen un equipamiento apropiado para obtener buenas imágenes, sumado a que no todos están dispuestos a ayudar y que lamentablemente la viabilidad de las muestras es bastante baja, casi provoca que este proyecto tuviese que detenerse en post de algún otro. Sin embargo, durante fines de octubre-principio de diciembre se pudo contar con la ayuda de 2 laboratorios, uno que aportó placas de muestras médicas y otro que además de muestras, contaba con un microscopio que podía usarse para recuperar imágenes del examen de Inmunofluorescencia Indirecta para Anticuerpos Autoinmunes.

En 3 fechas distintas se pudo conseguir más de 700 imágenes de distintos tipos de muestra, de las cuales unas 200 correspondían a imágenes de muestras negativas, unas 200 de imágenes con Patrón Homogéneo, unas 200 con Patrón moteado y el resto con diversos tipos de patrones de Inmunofluorescencia que no fueron considerados en este proyecto, todas las cuales tenían unas dimensiones de 1280 x 960 pixeles.

**2. Clasificación visual de las imágenes**

Una vez obtenidas las imágenes se procedió a la clasificación visual de todas las imágenes para poder dejar las que eran representativas de los patrones a clasificar, además se procedió a la edición y a la aumentación del número de imágenes obtenidas, con lo cual se mejoraban los datos que se usarían para el resto del proyecto, esto tomó varios días hasta alcanzar un número de casi 45.000 imágenes de los 3 tipos a estudiar de un tamaño de 320 x 320 pixeles.

**3. Aprendizaje de modelos de clasificaciones múltiples**

Aunque durante el primer semestre se había desarrollado un modelo que servía para el reconocimiento de 2 tipos de muestras, el reconocimiento representó un desafío adicional, ya que el manejo de etiquetas, número de imágenes y tiempo de procesamiento es distinto a lo ya realizado.

**4, 5 y 6. Creación de algoritmos de Clasificación, Entrenamiento de las Redes Neuronales y Ajustes de modelos y algoritmos.**

Estos 3 puntos se consideran juntos durante el desarrollo debido que a partir del primer modelado se debe ir realizando cambios, iteraciones ajustes pruebas y otras pruebas que incluyen tanto el tamaño, color y número de las imágenes con el fin de acercarse a niveles de Precisión y Pérdida adecuados, de hecho una de las modificaciones que hubo que realizar fue un ajuste al tamaño de las imágenes, reduciendo el tamaño de las imágenes de 320 x 320 pixeles (tamaño que es adecuado para observar las células en su totalidad) a un tamaño de 50 x 50 pixeles que es un valor más manejable por el hardware utilizado.

En una primera instancia se trató de desarrollar un modelo que abarcara los 3 tipos de resultados esperados, sin embargo se presentaron problemas tanto de los resultados de Precisión del modelo como de limitaciones del hardware, por lo que se optó por desarrollar 2 modelos que funcionarán en secuencia al momento de clasificar, uno que permite observar si una muestra es Negativa o no, y en caso de ser una muestra con algún tipo de positividad se aplicará el segundo modelo que es capaz de clasificar si una muestra tiene un Patrón Homogéneo o un Patrón Moteado. A continuación se presentan los resultados de ambos modelos, pero se dará énfasis en el segundo ya que es el que presenta mayor relevancia en el objetivo del proyecto.

El primer modelo generado constó de 17 capas y presentó una Precisión cercana a 96% tanto en el Entrenamiento como en la Validación del Modelo, pudiendo diferenciar las 100 muestras Negativas y las 200 muestras con patrones de positividad.

El segundo modelo constó de 14 capas de distintos tipos (Figura 17) y se realizó una iteración de 50 Epochs (Ciclos de Entrenamiento), para los cuales se utilizó el Set de casi 30.000 imágenes de ambos patrones (Homogéneo y Moteado con una proporción cercana al 1:1 de estos patrones) de los cuales 67%, escogido aleatoriamente, fue escogido para la Fase de Entrenamiento, mientras el 33% restante fue escogido para la realización de la Fase de Validación del modelo. En la Figura 18 se pueden observar los gráfico de Precisión tanto del Entrenamiento como de la Validación, llegando al final del ciclo 50 a una Precisión de un 93,72% para el Entrenamiento y de 91,02% para el caso de la Validación. Con respecto a la Pérdida del Modelo durante la fase de Entrenamiento alcanzó un 10%, mientras que en la fase de Validación llegó a ser un poco menor a 30% (Ver Figura 19). El Proceso por cada ciclo tomó entre 440 y 460 segundos, dando un tiempo final de casi 6 horas 40 minutos cada vez que se hacía una prueba, aunque no fuese exitosa.

|  |
| --- |
| Figura 17. Representación gráfica del Modelo desarrollado para la Clasificación de los Patrones Homogéneo y Moteado. |
|  |

|  |
| --- |
| Figura 18. Precisión del Modelo |
|  |

|  |
| --- |
| Figura 19. Pérdida del Modelo |
|  |

**7. Realización de Algoritmo de Evaluación de Imágenes**

De los modelos entrenados se obtuvieron los pesos y los modelos propiamente tal, los cuales fueron exportados a la Raspberry pi donde se procedió a generar un nuevo programa cuyo objetivo es la clasificación de nuevas muestras.

**8. Pruebas de Clasificación**

Las pruebas de clasificación se hicieron de manera individual, debido al bajo número de imágenes disponible, pudiendo ver que de las 40 imágenes elegidas al azar, reconoció correctamente todas ellas.

**Criterios de Aceptación del Sprint:**

- Crear un modelo con una precisión mayor al 90%.

- Reconocer imágenes individuales de cualquiera de los 3 posibles resultados.

**Problemas y desafíos:**

- Ajustar parámetros para no exceder las capacidades del Hardware.

- Crear un modelo universal que por sí mismo reconozca los 3 posibles resultados evaluados, y con una buena precisión

- Agregar más patrones al reconocimiento.

Durante el desarrollo del Sprint 3, uno de los principales problemas fue, además del problema de conseguir las imágenes fue el manejar los nuevos modelos junto a las limitaciones del Hardware. Estos puntos no se tomaron en consideración en el momento de planificar el Sprint 3 por lo que la estimación del tiempo fue muy corto para lo que realmente se necesitó (Ver Figura 20)

|  |
| --- |
| Figura 20. Burndown Chart del Sprint 3. |
|  |

# Capítulo V: Discusiones.

## 5.1 Discusiones.

Los requerimientos de Negocio y de Usuario y de Software se han podido trazar y se han resuelto, sin embargo los Requisitos de Negocio 4 y 5, y por tanto los requisitos de Usuario y de Software, que se trazan hasta ellos requerirán cambios posteriores ya que es posible que dependiendo de las características donde este sistema pueda ser implementados, requieran respaldar la información de las imágenes en sus propias bases de datos o que el envío de los resultados sea a un Sistema Informático de Laboratorio (LIS, por sus siglas en Inglés) que requiera algún módulo adaptador que actúe como puente e interprete entre ambos sistemas.

La Metodología SCRUM fue apropiada para un proyecto como este que buscaba innovar en un sistema que basándose en Deep Learning, permitiera apoyar a la labor diaria del personal de laboratorio. Sin embargo, existieron consideraciones como el tipo de tecnología usada y problemas que no se sopesaron en su momento, que llevó a que el proyecto se atrasara más allá de lo que se esperaba. Por otra parte el desarrollo Iterativo-Incremental con el que se trabajó este proyecto permitió trabajar sobre la base de las funcionalidades ya creadas en etapas anteriores lo que permitió obtener resultados de Entrenamiento y Validación de Modelos antes que de haber comenzado desde cero.

Por su parte el desarrollo de un sistema para el reconocimiento de imágenes de Inmunofluorescencia para Anticuerpos autoinmunes, ya había sido tratado por otros autores, Agrawal, P. y colaboradores (2013), desarrolló una revisión de diversas técnicas de clasificación que son de tipo matemático, mientras Wiliem, A. y colaboradores (2014), realizó una técnica de clasificación que definía unas figuras geométricas y cuyas características de brillo, tamaño, entre otros, determinaban el resultado si estas células pertenecían a tal o cual patrón, similar a esto fue un trabajo realizado por Siyamalan, M. y colaboradores (2015) quien utilizó una serie de técnicas complejas de distintos tipos para la clasificación de las imágenes obteniendo buenos resultados, siendo Gao, Z. y colaboradores (2015), quien utilizó por primera vez, con la bibliografía adquirida, Redes Neuronales Convolucionales para este propósito, incluso ellos probaron sus modelos en otros set de datos de características distintas a las que ellos usaron durante en el Entrenamiento, probando que los modelos creados por ellos eran dúctiles para su uso fuera de los medios controlados.

Respecto al uso de una CPU en lugar de una GPU compatible con CUDA, obliga a que las modelos sean unidimensionales provocando que sólo se pueda utilizar un tipo de capa neuronal por vez, lo que podría ir en desmedro de los buenos resultados del sistema, sin embargo los resultados obtenidos alcanzaron porcentajes de Precisión buenos. Empero, esta misma característica permitió poder utilizar los modelos en la Raspberry pi que es un sistema de características bastante básicas pero que clasificaba adecuadamente las muestras testeadas. El uso de Raspberry pi 3, en este caso, respondió a 2 temas, por una parte dar una solución al problema de que la tecnología de reconocimiento utiliza Linux y en los laboratorios el sistema operativo por excelencia es Windows, por lo que el tener un sistema que realice estas operaciones y que además funcione como Interface Web entre este sistema y el operador en una máquina con Windows, hacen a Raspberry pi un interesante prototipo, y por otra parte por un tema experimental, ya que se buscaba investigar si sus características la podría hacer apta para el desarrollo de equipos de laboratorios de bajo coste pero buenas prestaciones, ya la Raspberry pi 2, mostró sus potencialidades en proyectos de domótica (Cheuque, C., Baeza, F., Marquez, G., & Calderon, J. (2016)), así como para algunos proyectos de reconocimiento de imágenes (Kirpan, O. y colaboradores (2017)).

Con respecto al Riesgo técnico este en un momento pudo haber hecho fracasar por completo a este proyecto, debido a que, como se mencionó anteriormente, no se dispuso de imágenes apropiadas para ésta hasta finales de Octubre, y más aún fue en Noviembre cuando se pudo completar un volumen adecuado para el desarrollo completo de los modelos, en la Tabla 11 se grafica el impacto y la probabilidad de ocurrencia y el riesgo que significó cada uno de los puntos asignados al Riesgo Técnico donde el Impacto mientras más cercano a 1 implica que puede hacer fracasar por completo el proyecto, y en la Figura se observa cómo estos Riesgos Fueron Evolucionando.

|  |
| --- |
| Tabla 11. Matriz de Riesgo (Riesgo = Impacto \* Probabilidad de Ocurrencia) |
| |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  |  | Probabilidad de Ocurrencia | | | | | | Riesgo | Impacto | Agosto | Septiembre | Octubre | Noviembre | Diciembre | | Falta de Imágenes | 1 | 1 | 0,7 | 0,5 | 0 | 0 | | Pérdida de Información crítica o parte del código | 0,8 | 0,5 | 0,4 | 0,3 | 0,3 | 0,1 | | Falla del hardware donde operan los sistemas. | 0,3 | 0,2 | 0,2 | 0,2 | 0,2 | 0,2 | | Falta de experiencia en el desarrollo de alguno de los sistemas necesarios para el sistema. | 0,8 | 0,7 | 0,6 | 0,5 | 0,3 | 0,1 | | El uso de imágenes podría implicar aspectos legales que fueran necesarios de cubrir. | 0,4 | 0,3 | 0,2 | 0,1 | 0 | 0 | |

|  |
| --- |
| Figura 21. Evolución del Riesgo durante el Desarrollo del Proyecto |
|  |

Durante la Fase de desarrollo, y aunque los Sprint fueron planificados para durar 3 semanas aproximadamente, el desconocimiento de tecnologías y varios problemas de coordinación y otros aleatorios, llevaron a que cada Sprint durara al menos un 60% más de la planificación, lo que llevó a un atraso de todo el proyecto. Sin embargo se pudieron satisfacer casi completamente, todos los criterios de aceptación. En el Sprint 1 se logró dejar funcionando la API Rest Framework que ayuda a crear un proyecto más ordenado y con mayores medidas de seguridad, lo que es un punto destacable del proyecto. En el Sprint 2, fue el desarrollo de la Aplicación Web que se consiguió desarrollar una Web que funciona de acuerdo a lo que se planificaba en un principio, permitiendo un control de acceso dependiendo de quién ingresa, permitiendo aquí que cada usuario tenga la capacidad de ver y editar las peticiones ingresados por ellos mismos pudiendo observar las imágenes procesadas por el sistema junto a la clasificación para que así pueda decidir si de mejor forma el resultado a informar. Además esta Web le permite ingresar exámenes desde la misma plataforma para ser evaluados por el sistema. El último Sprint tardó más de lo esperado por lo limitado de las características del PC utilizado, llegando a realizar Entrenamientos que abarcaron hasta unas 25 horas, siendo la moda de cada Entrenamiento entre 4 a 7 horas, sin embargo logró resultados aceptables y permitió probar el reconocimiento de nuevas imágenes directamente en la Raspberry.

# Capítulo VI: Conclusiones

## 6.1 Conclusiones y trabajos posteriores

El desarrollo de sistemas que permitan la automatización de procesos hechos hasta ahora en forma manual, se está volviendo prioridad en el mundo actual donde las nuevas tecnologías permiten logros que hasta hace unas décadas eran impensables, esto es más patente aún en el caso de los exámenes realizados en el Laboratorio Clínico donde cualquier apoyo diagnóstico resulta en una disminución de los errores de origen humano y un beneficio a la salud del paciente. Considerando siempre que es el Personal Calificado quién finalmente será el responsable del informe médico.

El reconocimiento de los patrones de imágenes médicas mediante sistemas inteligentes, pese a tener sólo algunos años de investigación, avanza cada vez más rápido por lo que es cuestión de tiempo antes que aparezcan sistemas completos basados en esto principios.

Durante el desarrollo de este proyecto se logró obtener resultados más que interesantes (Precisión de 91% en la Validación de los Modelos) por lo que se deben realizar estudios más acabados donde profesionales de laboratorio puedan poner a prueba los resultados obtenidos por este sistema de apoyo, al tiempo de ayudar a mejorar la tecnología, cosa que es uno de los puntos fuertes de las técnicas de Deep Learning.

Por otra parte aunque los Procesos de Entrenamiento y Validación de modelos de Redes Neuronales Convolucionales resultan ser sumamente caros, en cuanto a recursos de CPU/GPU, RAM, Memoria física y Tiempo de trabajo, una vez realizado este proceso y con los datos generados guardados y exportados a otras plataformas, resulta notable los pocos recursos relativos que se requieren para poder evaluar nuevas imágenes, esto explica por qué el uso de una microcomputadora como Raspberry pi 3 resulta más que factible para el desarrollo de sistemas de bajo coste.

Como trabajo posterior quedará el llevar adelante el reconocimiento de Inmunofluorescencia para Anticuerpos Autoinmunes para que este modelo sea capaz de reconocer todo el universo de patrones existentes, y de ser probado en extenso en estudios de campo, llevando este proyecto a transformarse en una herramienta completa y confiable para su uso médico por parte del personal de laboratorio.

Por último este proyecto deja patente la utilidad de las Redes Neuronales Convolucionales como base para el reconocimiento de imágenes médicas, aportando un antecedente más a la bibliografía existente

# Bibliografía.

1. Arawal, P. y colaboradores. 2013. HEp-2 Cell Image Classification: A Comparative Analysis. Conference: International Workshop on Machine Learning in Medical Imaging.
2. Barrientos, R. y cols., 2009. Árboles de decisión como herramienta en el diagnóstico médico. Revista Médica de la Universidad Veracruzana, Julio - Diciembre, 2009, 19-24
3. BD Biosciences. BD FLUORESCENCE SPECTRUM VIEWER. Disponible en: <http://www.bdbiosciences.com/us/s/spectrumviewer> . Revisado: 05-10-2017.
4. Cheuque, C., Baeza, F., Marquez, G., & Calderon, J. 2016. Towards to responsive web services for smart home LED control with Raspberry Pi. A first approach. En Proceedings - 2015 34th International Conference of the Chilean Computer Science Society, SCCC 2015 (Vol. 2016-February).
5. Cires, D. y colaboradores. 2011.  Flexible, High Performance Convolutional Neural Networks for Image Classiﬁcation. Proceedings of the Twenty-Second international joint conference on Artificial Intelligence-Volume Volume Two 2: 1237–1242.
6. Gao, Z. y cols. 2015. HEp-2 Cell Image Classification with Deep Convolutional Neural Networks. Disponible en: <https://arxiv.org/abs/1504.02531>. Revisado: 20-10-2017.
7. IDEXX Laboratories, 2013. ELISA Technical Guide. Disponible en: <https://www.idexx.com/pdf/en_us/livestock-poultry/elisa-technical-guide.pdf>. Revisado: 23-04-2017.
8. Kirpan, O. y colaboradores. 2017. Object Detection on Raspberry Pi. International Journal of Engineering Science and Computing. Volumen 7 Issue No.3.
9. Koivunen, M., y Krogsrud, R., 2006. Principles of Immunochemical Techniques used in Clinical Laboratories. Lab Medicine. Volúmen 37, Número 8. Págs.: 490-497.
10. LeCun, Y. y colaboradores, 2015. Deep learning. Nature. Vol: 521.
11. Litjens, G. y cols., 2017. A Survey on Deep Learning in Medical Image Analysis.Disponible: https://arxiv.org/abs/1702.05747. Revisado: 25-07-2017.
12. Lowd, D. y Domingos, P., 2005. Naive Bayes Models for Probability Estimation. Proceedings of the 22nd International Conference on Machine Learning, Bonn, Germany, 2005.
13. Micucci, H. y cols. 2010. 50 años en el Laboratorio Clínico. *Acta bioquímica clínica latinoamericana*, *44*(4), 711-722. Recuperado en 09 de mayo de 2017, de <http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S0325-29572010000400013&lng=es&tlng=es>.
14. Palomo I., y cols., 2002. Fundamentos de Inmunología Básica y Clínica. Editorial Universidad De Talca. Primera Edición.
15. Rosenblum, M., Remedios, K., y Abbas A. 2015. Mechanisms of human autoimmunity. The Journal of Clinical Investigation. N° 125(6). Págs:2228-2233.
16. Schmidhuber, J., 2008. Deep Learning in Neural Networks: An Overview. Technical Report IDSIA-03-14 / arXiv:1404.782.
17. Siyamalan, M. y colaboradores. 2015. An automated pattern recognition system for classifying indirect immunofluorescence images for HEp-2 cells and specimens. Pattern Recognition, 51, 12-26.
18. Vircell Microbiologists. Galería de Imágenes. En: <http://www.vircell.com/soporte-tecnico/galeria-de-imagenes/DFA/> . Revisado: 05-10-2017.
19. Wiliem, A. y colaboradores. 2014. Automatic Classification of Human Epithelial Type 2 Cell Indirect Immunofluorescence Images using Cell Pyramid Matching. Disponible en: <https://arxiv.org/abs/1403.3780>. Revisado: 01-12-2017.

1. [↑](#footnote-ref-1)