

Mar del Plata, 27 de abril de 2022.-

## RESOLUCIÓN DEL RECTORADO N° 262/22

### VISTO:

El proyecto de investigación “*Inteligencia Artificial en FPGA. IA-FPGA – Etapa II*” presentado por los docentes de la Facultad de Ingeniería, Dr. Martín Osvaldo Vázquez, Ing. Bruno Eduardo Nicolás Constanzo, Mg. Ing. Elías Lucas Leiva y Mg. Jordina Torrens Barrena tramitado bajo el expediente de Investigación 265-2022; y

### CONSIDERANDO:

La presentación formal del proyecto se hizo de acuerdo a lo establecido en el art. 31 de la Resolución de Rectorado N° 053/18 mediante expediente abierto a solicitud de la Secretaria de Investigación de la Facultad de Ingeniería, con fecha 20 de abril de 2022;

La evaluación del especialista Dr. Juan M. Toloza, y del Dr. Ruben S. Wainschenker;

Que la Secretaria de Investigación de la Universidad FASTA, la Secretaria de Investigación de la Facultad de Ingeniería y el Decano de la Facultad de Ingeniería, han propuesto al Dr. Martín Osvaldo Vázquez como Director e Investigador Titular categorizado por la UFASTA;

Que el investigador propuesto acredita una dilatada y rica trayectoria científica, profesional y académica que permite destacar su testimonio y vocación de servicio y que cualifica indudablemente el claustro universitario;

Que la Secretaria de Investigación de la Universidad FASTA, la Secretaria de Investigación de la Facultad de Ingeniería y el Decano de la Facultad de Ingeniería, han propuesto al Ing. Bruno Eduardo Nicolás Constanzo como Investigador Adjunto categorizado por la UFASTA;

Que el investigador propuesto acredita una dilatada y rica trayectoria científica, profesional y académica que permite destacar su testimonio y vocación de servicio y que cualifica indudablemente el claustro universitario;

Que la Secretaria de Investigación de la Universidad FASTA, la Secretaria de Investigación de la Facultad de Ingeniería y el Decano de la Facultad de Ingeniería, han propuesto como auxiliares de investigación al Mg. Ing. Elías Lucas Leiva y la Mg. Jordina Torrens Barrena;

Lo dispuesto por la Resolución del Rectorado N° 053/18, en sus artículos 8, 9, 10, 17 y concordantes;

Por ello, y en uso de las atribuciones que le confieren los Arts. 28° inc. d) y concordantes del Estatuto Universitario

**EL RECTOR DE LA UNIVERSIDAD FASTA  
DE LA FRATERNIDAD DE AGRUPACIONES SANTO TOMÁS DE AQUINO  
RESUELVE :**

**Artículo 1º:** Aprobar el proyecto de investigación “*Inteligencia Artificial en FPGA. IA-FPGA – Etapa II*” que se adjunta en Anexo I de a presente.-


**Artículo 2º:** Designar al **Dr. Martín Osvaldo VÁZQUEZ** (DNI 23.617.506) como Director e Investigador Titular categorizado por la UFASTA del Proyecto “*Inteligencia Artificial en FPGA. IA-FPGA – Etapa II*” por el término de 30 meses a partir del 1º de mayo de 2022, según los alcances del art. 32 y cc. de la Resolución Rectoral N° 053/18.-

**Artículo 3º:** Designar al **Ing. Bruno Eduardo Nicolás CONSTANZO** (DNI 33.566.742) como Investigador Adjunto categorizado por la UFASTA del Proyecto “*Inteligencia Artificial en FPGA. IA-FPGA – Etapa II*” por el término de 30 meses a partir del 1º de mayo de 2022.-

**Artículo 4º:** Designar al **Mg. Ing. Elías Lucas LEIVA** (DNI 27.539.740) como Auxiliar de Investigación Graduado categorizado por la UFASTA del Proyecto “*Inteligencia Artificial en FPGA. IA-FPGA – Etapa II*” por el término de 30 meses a partir del 1º de mayo de 2022.-

**Artículo 5º:** Designar a la **Mg. Jordina TORRENS BARRENA** (DNI 48.014.960-E) como Auxiliar de Investigación Graduado categorizado por la UFASTA del Proyecto “*Inteligencia Artificial en FPGA. IA-FPGA – Etapa II*” por el término de 30 meses a partir del 1º de mayo de 2022.-

**Artículo 6º** Dése a conocer, remítase copia a la Secretaría de Investigación de la Universidad FASTA, al Decanato de la Facultad de Ingeniería y a los investigadores designados, archívese.



PROF. MARCELA S. GRECA DE GIACAGLIA  
SECRETARÍA GENERAL  
UNIVERSIDAD FASTA



DR. JUAN CARLOS MENA  
RECTOR  
UNIVERSIDAD FASTA

ANEXO

RESOLUCIÓN DEL RECTORADO N° 262/22

Proyecto de Investigación “*Inteligencia Artificial en FPGA. IA-FPGA – Etapa II*”

1. EL PROYECTO DE INVESTIGACIÓN

TÍTULO O DENOMINACIÓN DEL PROYECTO

INTELIGENCIA ARTIFICIAL EN FFGA – ETAPA 2 (IA-FPGA-E2)

MES Y AÑO DE INICIO: 05 / 2022

MES Y AÑO DE FINALIZACIÓN: 10 / 2024

LÍNEA DE INVESTIGACIÓN: Sistemas Embebidos

GRUPO DE INVESTIGACIÓN: Sistemas Embebidos

ÁREA DE CONOCIMIENTO: INGENIERIAS Y TECNOLOGÍAS

TIPO DE INVESTIGACIÓN: APLICADA

2. INSTITUCIONES PARTICIPANTES

INSTITUCIÓN/ES COLABORADORA/S DEL PROYECTO:

3. DIRECTOR

NOMBRE Y APELLIDO DEL DIRECTOR DEL PROYECTO: Dr. Martín Vázquez

DIRECCIÓN DE CONTACTO DEL DIRECTOR (telefónica y/o electrónica): [movazquez@ufasta.edu.ar](mailto:movazquez@ufasta.edu.ar)

NOMBRE Y APELLIDO DEL CO-DIRECTOR (si lo/s hubiera): Ing. Bruno Constanzo

DIRECCIÓN DE CONTACTO DEL CO-DIRECTOR (telefónica y/o electrónica) (si lo/s hubiera):  
[bconstanzo@ufasta.edu.ar](mailto:bconstanzo@ufasta.edu.ar)



#### 4. EQUIPO DE TRABAJO

NOMBRE Y APELLIDO	UNIDAD ACADÉMICA / INSTITUCIÓN	CATEGORÍA / FUNCIÓN	HORAS DESIGNACIÓN
DR. MARTÍN VAZQUEZ	FACULTAD DE INGENIERÍA	INVESTIGADOR TITULAR/DIRECTOR	20
ING. BRUNO CONSTANZO	FACULTAD DE INGENIERÍA	INVENTIGADOR ADJUNTO/CODIRECTOR	5
MG. LUCAS LEIVA	FACULTAD DE INGENIERÍA	AUXILIAR INVESTIGADOR GRADUADO/INTEGRANTE	20
DRA. JORDINA TORRENTS-BARRENA	FACULTAD DE INGENIERÍA	AUXILIAR INVESTIGADOR GRADUADO/INTEGRANTE	2

#### 5. CARACTERIZACIÓN DEL PROYECTO

##### **Descripción breve del proyecto :**

Se trata de un proyecto de investigación que continúa con la investigación y desarrollo realizados en un proyecto precedente "IA-FPGA" ejecutado en la misma unidad académica. El proyecto apunta al desarrollo e implementación de algoritmos de Inteligencia Artificial, en particular técnicas de *Deep Learning* (DL) basadas en Redes Neuronales Convolutivas (CNN), impartiendo especial énfasis a la aceleración de soluciones existentes a problemas particulares. Las exigencias actuales de los usuarios requieren la utilización de plataformas de procesamiento no convencionales para satisfacer las expectativas. Una de las plataformas que se utiliza para este fin son las basadas en lógica programable *FPGAs* (*Field Programmable Gate Array*) brindando además soporte para el procesamiento embebido, portable y móvil, a través del uso de dispositivos SoC (*System-on-a-Chip*) que combinan la capacidad de cómputo secuencial (procesador) y paralelo (FPGA).

Durante la ejecución del proyecto precedente se desarrollaron dos líneas de trabajo. Una línea consistió en la aceleración hardware basado en FPGA y en HLS (*High Level Synthesis*), de problemas que requieren computación intensiva, tales casos de predictores LibSVM y *Reinforcement Learning* (RL). En lo que respecta a la otra línea, se efectuó relevamiento vinculado al desarrollo de inteligencia Artificial (IA) en plataformas embebidas basadas en FPGAs, selección de tecnología e implementación de un caso de estudio simple de trabajo.

En este proyecto (etapa 2) se formalizará un método general automático o semi-automático de aceleración hardware basado en FPGA, de problemas que requieren computación intensiva en el dominio de Inteligencia Artificial mediante la utilización de HLS. Además, se diseñarán e implementarán casos novedosos y desafiantes de sistemas embebidos basados en lógica programable mediante el uso de DL.

**Problema o necesidad a resolver:**

El modelo de Aprendizaje Profundo de redes CNN ha empujado los límites de varias tareas de Inteligencia Artificial. En algoritmos típicos de *Machine Learning* (ML), las CNN involucran dos fases: una fase de entrenamiento y otra fase de inferencia o clasificación. En los últimos años han surgido diversos *frameworks* que favorecen el diseño e implementación de sistemas de Inteligencia Artificial, como PyTorch, TensorFlow, Lasagne basado en Theano, Keras, y Chainer, entre otros. Estos *frameworks* proporcionan unidades de red neuronal, funciones de costo y optimizadores para ensamblar y entrenar modelos.

Se aprecian dos aspectos fundamentales a considerar en cuanto al desarrollo involucrado en CNN. Por un lado, la aceleración hardware del entrenamiento de la red. Por otro lado, en el caso que se requiera, la implementación de una solución final completa en un sistema embebido, el cuál debe considerar restricciones duras de tiempo real, ocupación de área, consumo de potencia energética, etc.

El empleo de la tecnología FPGA aparece como una buena alternativa en el momento de utilizarse como plataforma para la implementación de arquitecturas de redes neuronales complejas, o para ser utilizadas como co-procesadores hardware dedicados, ejecutando las rutinas críticas de los sistemas de inteligencia artificial que dominan actualmente el mercado.

**Resumen, detallando objetivos generales y particulares:**

El objetivo general es la implementación en el ámbito Universitario de actividades de desarrollo tecnológico de punta, que promueva el desarrollo en Argentina de industrias vinculadas con la electrónica y el diseño de sistemas embebidos basados en lógica programable y.

Otro de los objetivos principales es la investigación, aplicación y mejoras de plataformas existentes basadas en FPGA para el desarrollo de soluciones embebidas que utilizan DL.

Por otra parte, se espera desarrollar métodos automáticos o semi-automáticos de aceleración hardware basado en FPGA y HLS, de algoritmos de ML que requieran computación intensiva.

Como objetivos particulares se encuentra el diseño e implementación de:

- Plataformas de Inteligencia Artificial basadas en FPGAs de Xilinx e Intel FPGA.
- Co-procesadores hardware basados en lógica programable, que aceleren las técnicas de Inteligencia Artificial basadas en DL, tanto en las fases de entrenamiento, como en clasificación/inferencia.
- Arquitecturas de CNN aplicadas a casos de estudio particulares.
- Generalización de las arquitecturas e integración con *frameworks* de DL.

**Actividades del proyecto:**

Durante la ejecución del proyecto se implementarán arquitecturas hardware que permitan la aceleración de técnicas de inteligencia artificial, principalmente de DL basado en CNN. Dentro de las características se pretende:

- Obtener tiempos de ejecuciones significativamente inferiores a alternativas software (10-20x)
- Permitir la ejecución en tiempo real
- Implementar precisiones parametrizables, a medida de la aplicación

- Alcanzar un ahorro en el consumo de energía
- Permitir la ejecución en sistemas embebidos y/o portables.
- Aplicar optimizaciones que se adecúen a la plataforma destino.

El proyecto comenzará con la profundización en el perfeccionamiento de los integrantes del grupo en las tecnologías de síntesis en los ambientes de diseño provistos por los fabricantes de la tecnología objetivo.

Luego se profundizará el estudio de las técnicas de implementación en DL, particularmente implementación hardware de redes CNN. En este período se planea profundizar los estudios realizados en proyectos anteriores, y fortalecer la vinculación con otros institutos de investigación

Se ahondará en el análisis de los algoritmos utilizados en DL y se aplicarán las arquitecturas generadas a casos de estudios particulares.

A continuación, se generalizarán de las arquitecturas implementadas e integrarán con *frameworks* de DL.

Por último, los detalles de las implementaciones se consignarán en informes y eventualmente publicarán y/o presentarán en revistas y congresos afines.

#### **Novedad u originalidad en el conocimiento:**

La Inteligencia Artificial (IA) ha tenido, desde hace algunos años, un gran impacto en la sociedad y en la economía, entre otras áreas. Hoy en día muchas organizaciones utilizan estas herramientas para apalancar su crecimiento. Según algunos reportes de consultoras, la utilización de IA podría resultar en un incremento de más del 35% de producto bruto interno de algunos países.

En el contexto nacional, se cuentan con varias iniciativas orientadas a mejorar la formación de recursos humanos en la informática, como son: la reciente Ley De Economía del Conocimiento<sup>1</sup> para aumentar la productividad y la generación de empleo de calidad, que incluye a las actividades relacionadas con el software; y el plan industria 4.0<sup>2</sup> para favorecer la competitividad del sector productivo a través de la incorporación de tecnologías emergentes como la IA, entre otras. En particular se desea mencionar el Plan Nacional de Inteligencia Artificial, que es un plan estratégico que se encuentra en la Agenda Digital 2030<sup>3</sup>.

Por otro lado, en los países en vía de desarrollo como es el caso de Argentina, la utilización de tecnologías como FPGAs para el desarrollo de hardware se presenta como una alternativa muy auspiciosa. Las razones fundamentales se encuentran asociadas a los costos en comparación con desarrollos ASICs, tanto en lo que respecta al chip adquirido como también al diseño de estos

<sup>1</sup> <https://www.argentina.gob.ar/noticias/el-congreso-aprobo-la-ley-que-promueve-la-economia-del-conocimiento>

<sup>2</sup> <https://www.argentina.gob.ar/noticias/el-gobierno-creo-el-plan-industria-argentina-40-para-favorecer-la-transformacion>

<sup>3</sup> <https://www.casariosada.gob.ar/informacion/actividad-oficial/9-noticias/44081-el-gobierno-presento-la-nueva-agenda-digital-2030>

circuitos. Los FPGAs son una tecnología accesible que ofrece la posibilidad de realizar implementaciones a partir de descripciones basadas en software.

En el ámbito académico argentino no se han desarrollado demasiados proyectos sobre sistemas embebidos que implementen algoritmos de DL. Por otro lado, la aplicación de la tecnología FPGA para el desarrollo de coprocesadores aceleradores y/o sistemas embebidos en el dominio de DL, se encuentra en un estado muy insipiente a nivel nacional, con lo cual presenta una gran oportunidad para realizar contribuciones y ofrecer soluciones innovadoras.

#### Resultados esperados

- Promoción y aplicación de la tecnología ML/DL en sistemas embebidos basados en lógica programable en la región.
- Desarrollo de métodos automáticos, o semi-automáticos, para de aceleración hardware mediante el uso coprocesador basados FPGA, de algoritmos que requieren computación intensiva en aplicaciones de ML/DL. La aceleración se abordará tanto para la fase de entrenamiento como para la fase de predicción o clasificación.
- Diseño e implementación de casos novedosos y desafiantes de sistemas embebidos basados en FPGA mediante el uso de ML y DL. Se apuntará a aplicaciones que tengan que ver con la industria 4.0, tales casos Medicina 4.0, Agro 4.0, entre otros.

#### Impacto de los resultados:

- Culminación de tesis doctorales de integrantes del proyecto
- Publicaciones en revistas internacionales con índice de impacto. Se dirigirá fundamentalmente hacia el índice JCR (*Journal Citation Report*), pero también hacia índices menores como *LatinIndex*, entre otros.
- Publicaciones en congresos internacionales.
- Publicaciones y presentaciones en congresos nacionales. El objetivo de difundir y promover el trabajo realizado en el ámbito local

#### Interés para la Universidad FASTA :

En la UFASTA este proyecto es innovador y apunta a sensibilizar los recursos humanos locales acerca de la apertura en la Argentina en los mercados de productos vinculados a la aceleración de aplicaciones cómputo intensivo en el dominio de ML y DL, a los sistemas embebidos y a los sistemas digitales. Se espera que el conocimiento generado durante la ejecución del proyecto, permita generalizar una plataforma de soporte de inteligencia artificial (particularmente CNN) en hardware aplicable en diferentes campos, resultando un producto conveniente a otros investigadores y empresas.





Uno de los principales intereses de este proyecto es que la UFASTA contribuya en la región en el desarrollo de industrias vinculadas con la electrónica y el diseño de sistemas digitales, en las principales tendencias globales de investigación. El proyecto pretende formar recursos locales en tecnología de punta y además vincularse con otros centros de investigación de la región, como es el caso de Universidad de Mar del Plata y Universidad Nacional del Centro de la Pcia. Bs.As. Se espera que estos vínculos se sustancien mediante la ejecución de tesinas de grados, PPS (prácticas profesionales supervisadas) y eventualmente trabajos de postgrados. Además se pretende realizar publicaciones internacionales con índice de impacto y difusión en congresos nacionales afines a la temática del proyecto.

## 6. DISEÑO METODOLÓGICO

En proyectos anteriores ejecutados por el grupo de investigación, se adquirieron tarjetas de desarrollo de Xilinx e Intel FPGA. Los integrantes del grupo ya tienen conocimientos teóricos y prácticos acerca de estos ambientes de desarrollo y en lo que se refiere a lenguajes de descripción de hardware (HDL), herramientas de síntesis de alto nivel y prácticas con las herramientas de diseño en FPGA. Además, miembros del proyecto poseen dilatada experiencia en el campo de ML.

Durante la ejecución del proyecto anterior “Inteligencia Artificial en FPGA (IA-FPGA)”, se desarrollaron dos líneas de trabajo. Una línea (línea 1) consistió en la aceleración hardware basado en FPGA de predictores LibSVM, y además con la exploración en el área de *Reinforcement Learning* (RL). En cuanto a la otra línea (línea 2), se efectuó relevamiento vinculado al desarrollo de inteligencia Artificial (IA) en plataformas embebidas basadas en FPGAs, selección de tecnología e implementación de un caso de estudio inicial de trabajo. Este proyecto será una continuación (Etapa 2) del proyecto que finaliza. En esta etapa se profundizará con las dos líneas de trabajo desarrolladas en el proyecto anterior:

- En la **línea 1**, el proyecto anterior se introduce un método de aceleración de la librería LibSVM usando co-procesadores FPGA mediante el uso de síntesis de alto nivel. La librería LibSVM posee rutinas que implementan el estándar SVM (*Support Vector Machine*), el cuál es una técnica de ML muy usada por la comunidad científica para clasificadores supervisados. Los clasificadores SVM poseen dos etapas: entrenamiento y predicción. La librería LibSVM es una implementación muy utilizada por sus prestaciones en cuanto a tiempos de cómputo y precisión, así como también por la característica que sus códigos se encuentran disponibles y abiertos (*open source*). Los pilares fundamentales en las que se sostiene la metodología propuesta, se encuentran vinculados a la síntesis de alto nivel basada en descripciones C o C++, al uso de dispositivos flexibles en cuanto a reconfiguración como FPGAs, y a las herramientas HLS disponibles, las cuales mediante el uso adecuado de directivas obtienen descripciones hardware que posibilitan la implementación eficiente en el dispositivo FPGA.





Otro dominio que se abordó vinculado a esta línea de investigación es el *Reinforcement Learning*. El RL es un área de ML en donde se entrena una entidad (agente) para realizar una tarea determinada. A diferencia de las técnicas clásicas, no requiere de fases de entrenamiento e inferencia separadas, ya que se basa en un enfoque de prueba y error. Por sus características, RL es una herramienta poderosa para resolver problemas donde el escenario es desconocido o dinámico en el tiempo. Generalmente este tipo de aplicaciones requieren de potencia de cómputo para el procesamiento de grandes volúmenes de datos y un consumo de energía limitado. Se desarrolló un acelerador de RL para el Proceso de Decisión de Markov implementado en la librería pública IA-Toolbox, mediante la aplicación de la metodología propuesta basada en herramientas de síntesis de alto nivel.

**En la Etapa 2, para la línea 1 se formalizarán métodos generales automáticos o semi-automáticos de aceleración hardware basado en FPGA, de problemas que requieren computación intensiva en el dominio de ML mediante la utilización de HLS. Los métodos propuestos servirán para abordar la aceleración en algoritmos o librerías de ML, utilizados tanto en el entrenamiento como en la clasificación o predicción.**

- En lo que respecta a la **línea 2**, en el proyecto anterior primero se hizo relevamiento de la de la tecnología y las herramientas para el desarrollo de sistemas de inteligencia artificial en FPGAs. De este relevamiento, en una primera fase surgieron diferentes opciones: el desarrollo de una arquitectura propietaria mediante, el uso de la plataforma PYNQ, el uso de DNNDK y de Vitis-AI. En la segunda fase del proyecto finalmente se optó por el uso de PYNQ y Vitis-AI. Vitis-AI es un entorno de desarrollo que contiene núcleos IP optimizados, herramientas, bibliotecas, modelos y diseños de ejemplo, brindando una eficiencia y facilidad de uso para la aceleración de IA en FPGA de Xilinx. Este entorno se basa en el uso de DPU (*Deep-Learning Processor Unit*) el cual es un motor de computación configurable optimizado para redes neuronales convolutivas. Por otro lado, PYNQ es un proyecto de código abierto de Xilinx, que integra la tecnología con Python, siendo su principal objetivo la facilitación del diseño de aplicaciones a desarrolladores. Conceptualmente opera a través del uso *overlays*, que son configuradas dinámicamente en la lógica programable, y utilizados desde un programa Python ejecutado desde un Jupyter Notebook en el sistema de procesamiento (PS) del FPGA. Uno de estos *overlays* disponible es DPU-PYNQ, que contiene ejemplos de inferencia y entrenamiento para ejecutarse sobre plataformas PYNQ. Este *overlay* integra PYNQ con Vitis-AI simplificando el proceso de diseño de aplicaciones basadas en DL.

Entre los *frameworks* de DL disponibles en la actualidad, se destaca TensorFlow. Este *framework* soporta el *deploy* de modelos entrenados en diversos formatos y plataformas, incluyendo a Keras. Keras brinda abstracciones que permiten generar *pipelines* de entrenamiento de modelos de DL de gran complejidad de manera simple.

En el proyecto anterior se estableció un método para el desarrollo de aplicaciones IA embebidas en FPGAs, en donde el entrenamiento de los modelos DL se separa del entorno de programación de Vitis-AI. Esto permite a los expertos en DL trabajar en un entorno familiar y luego exportar y procesar (con el *toolchain* de Vitis-AI) los modelos en el formato adecuado, para generar la implementación de la inferencia embebida en el FPGA. Como caso de estudio se trabajó con el conjunto de dígitos manuscritos MNIST, utilizando los primeros 10,000 datos para entrenamiento. Se utilizó como optimizador el algoritmo Adam. Como plataforma para



implementación se utilizó el kit de desarrollo Ultra96V2 de Avnet, que cuenta con un MPSoC (*Multi-Processor-on-a-Chip*), por ser la plataforma más económica compatible con las herramientas de desarrollo.

**En la Etapa 2, para la línea 2 se espera profundizar todos los conocimientos adquiridos en esta línea en el proyecto anterior. Se diseñarán e implementarán casos novedosos y desafiantes de sistemas embebidos en FPGAs mediante el uso de DL. Se intentará dirigir a aplicaciones de Industria 4.0. Uno de los casos iniciales que se comenzará abordar es el conteo automático de huevos de parásitos en heces de ganado, mediante la utilización de la plataforma y tecnologías seleccionadas en la ejecución de la línea del presente proyecto, a través de la vinculación con el Centro de Investigación Veterinaria Tandil (CIVETAN).**

Durante la ejecución del proyecto se estimulará la comunicación constante entre los diferentes integrantes mediante ponencias y realización de reportes técnicos internos. Estos reportes son muy importantes para la comunicación científica, formalización de resultados y para establecer el paso inicial a publicaciones en revistas y congresos, así como también en capítulos de tesis doctorales.

Además se desarrollarán las actividades en el laboratorio de investigación perteneciente Universidad de FASTA, pero algunas tareas se llevarán a cabo en varias universidades asociadas tales como la Universidad de Mar del Plata, la Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), Universidad Autónoma de Madrid (UAM). Además, el proyecto se vinculará con el Centro de Investigación Veterinaria Tandil (CIVETAN) y el Laboratorio de Investigación y Desarrollo de Tecnologías en Informática Forense de Facultad de Ingeniería de Universidad de FASTA (InfoLab).

## 7. BIBLIOGRAFÍA

[1] Xilinx, "Vitis AI User Guide". Disponible en <https://docs.xilinx.com/r/1.3-English/ug1414-vitis-ai/Revision-History>, 2022.

[2] "PYNQ : "Python Productivity". Disponible en <http://www.pynq.io/>, 2022.

[3] Xilinx, "PYNQ-DPU : DPU on PYNQ". Disponible en <https://github.com/Xilinx/DPU-PYNQ>.

[4] W. Zhu, L. Xie, J. Han and X.Guo. "The application of deep learning in cancer prognosis prediction". *Cancers*, vol. 12 no 3, p. 603, 2020.

[5] S. Minaee, Y.Y. Boykov, F. Porikli, A.J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey". *IEEE Transaction on Pattern Analysis and Machine Intelligent*, 2021

[6] Y. Wang, X. Zhu, B. Wu. "Automatic detection of individual oil palm trees from UAV images using HOG features and an SVM classifier". *International Journal of Remote Sensing*. 40, 2019

- [7] H. Younes, A. Ibrahim, M. Rizk and M. Valle. "Algorithmic-level approximate tensorial SVM using High-Level Synthesis on FPGA". *Electronics*, 10, 2021
- [8] C. Kardaris. "Acceleration of SVM machine Learning algorithm on reconfigurable FPGA logic Plaforms". *Ntional Technicals University of Athens*, 2020.
- [9] Kehtarnavaz, Nasser, and Mark Gamadia. "Real-time image and video processing: from research to reality." *Synthesis Lectures on Image, Video & Multimedia Processing 2.1*: 1-108, 2006.
- [10] Automated Imaging Association, <http://www.visiononline.org/>
- [11] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," in *NIPS Workshop Autodiff*, 2017.
- [12] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. Software available from tensorflow.org.
- [13] R. Al-Rfou, et. al, "Theano: A python framework for fast computation of mathematical expressions," *arXiv preprint arXiv:1605.02688*, 2016.
- [14] S. Dieleman, J. Schlüter, C. Raffel, E. Olson, S. K. Sønderby, D. Nouri, D. Maturana, M. Thoma, E. Battenberg, J. Kelly, J. D. Fauw, M. Heilman, D. M. de Almeida, B. McFee, H. Weideman, G. Takács, P. de Rivaz, J. Crall, G. Sanders, K. Rasul, C. Liu, G. French, and J. Degraeve, "Lasagne: First release.," Aug. 2015.
- [15] F. Chollet et al., "Keras." <https://keras.io>, 2015.
- [16] S. Tokui, K. Oono, S. Hido, and J. Clayton, "Chainer: a next-generation open source framework for deep learning," in *NIPS Workshop Machine Learning Systems*, 2015.
- [17] Rodríguez-Andina, Juan J., Maria D. Valdes-Pena, and Maria J. Moure. "Advanced features and industrial applications of FPGAs—A review." *IEEE Transactions on Industrial Informatics* 11.4 (2015): 853-864.
- [18] Yann LeCun, Yoshua Bengio, and Georey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.
- [19] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, and others. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015.
- [20] Ross Girshick. Fast R-CNN. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition- CVPR '15*, pages 1440–1448, 2015.
- [21] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully Convolutional Networks for Semantic Segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition - CVPR '15*, pages 3431–3440, 2015.

- [22] Ying Zhang, Mohammad Pezeshki, Philémon Brakel, Saizheng Zhang, Cesar Laurent Yoshua Bengio, and Aaron Courville. Towards end-to-end speech recognition with deep convolutional neural networks. arXiv preprint, arXiv:1701, 2017.
- [23] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint, arXiv:1409:1–14, 2014.
- [24] Eriko Nurvitadhi, Suchit Subhaschandra, Guy Boudoukh, Ganesh Venkatesh, Jaewoong Sim, Debbie Marr, Randy Huang, Jason OngGeeHock, Yeong Tat Liew, Krishnan Srivatsan, and Duncan Moss. Can FPGAs Beat GPUs in Accelerating Next-Generation Deep Neural Networks? In Proceedings of the ACM/SIGDA International Symposium on Field-Programmable Gate Arrays - FPGA '17, pages 5–14, 2017.
- [25] Kalin Ovtcharov, Olatunji Ruwase, Joo-young Kim, Jeremy Fowers, Karin Strauss, and Eric Chung. Accelerating Deep Convolutional Neural Networks Using Specialized Hardware. White paper, pages 3–6, 2, 2015.
- [26] Jiantao Qiu, Jie Wang, Song Yao, Kaiyuan Guo, Boxun Li, Erjin Zhou, Jincheng Yu, Tianqi Tang, Ningyi Xu, Sen Song, Yu Wang, and Huazhong Yang. Going Deeper with Embedded FPGA Platform for Convolutional Neural Network. In Proceedings of the ACM/SIGDA International Symposium on Field-Programmable Gate Arrays - FPGA '16, pages 26–35, New York, NY, USA, 2016. ACM.
- [27] Intel FPGA. Intel® Stratix® 10 Variable Precision DSP Blocks User Guide. Technical report, Intel FPGA Group, 2017.

