PROGRAMA DE CURSO

Código	Nombre		
CC7221	Reconocimiento Visual con Deep Learning		
Nombre en Inglés			
Visual Recognition with Deep Learning			
Créditos	**Horas de Cátedra**	**Horas Docencia Auxiliar**	**Horas de Trabajo Personal**
6	3	1.5	5,5
Requisitos		**Carácter del Curso**	
- CC3001 (Obligatorio) - Haber tomado alguno de los siguientes ramos: **CC5114, CC5206, CC6204, CC5509, EL4106, EL5206, EL7006**		Electivo para pre y postgrado en Computación.	
Competencias a las que tributa el curso			
Competencias Específicas:			
CE1:	Analizar problemas computacionales, construir modelos, expresándolos en representaciones y lenguajes formales adecuados.		
CE2:	Analizar, diseñar y/o adaptar algoritmos y estructuras de datos que cumplan con las garantías requeridas de correctitud y eficiencia.		
CE4:	Extraer información relevante, utilizando el proceso de descubrimiento de conocimiento de datos.		
CE9:	Desarrollar soluciones computacionales de manera interdisciplinaria y colaborativa.		
Competencias genéricas			
CG1:	Comunicación profesional y académica: Comunicar en español de forma estratégica, clara y eficaz, tanto en modalidad oral como escrita, puntos de vista, propuestas de proyectos y resultados de investigación fundamentados, en situaciones de comunicación compleja, en ambientes sociales, académicos y profesionales.		
CG2:	Comunicación en inglés: Leer y escuchar de manera comprensiva en inglés variados tipos de textos e informaciones sobre temas concretos o abstractos, comunicando experiencias y opiniones, adecuándose a diferentes contextos de acuerdo a las características de la audiencia.		
CG3:	Compromiso ético: Actuar de manera responsable y honesta, dando cuenta en forma crítica de sus propias acciones y sus consecuencias, en el marco del respeto hacia la dignidad de las personas y el cuidado del medio social, cultural y natural.		

CG5:	Innovación: Concebir ideas viables y novedosas para resolver problemas o necesidades, materializadas en productos, servicios o en mejoras a procesos, considerando el contexto sociocultural, económico y los beneficios para el usuario.
CG6:	Emprendimiento: Identificar oportunidades aprovechables para crear e introducir nuevos servicios o productos con valor económico y social, a partir de la toma de decisiones en un contexto complejo de incertidumbre, demostrando compromiso e iniciativa en su quehacer.

Propósito del curso
En los últimos años hemos visto significativos avances en el área de Inteligencia Artificial, particularmente sobre la base del enfoque basado en Deep Learning. Este hecho ha permitido el desarrollo de una gran cantidad de aplicaciones en diversos campos industriales. Por otro lado, también hemos visto una explosión en la generación de datos debido a la proliferación de dispositivos móviles, así como a la masificación de las redes sociales. Uno de estos datos son las imágenes, sobre las que se hace necesario aplicar métodos automáticos para interpretarlas. *En este curso, analizaremos el estado del arte en reconocimiento visual bajo el enfoque de deep learning. Al término del ramo, el alumno será capaz de entender y analizar modelos en clasificación, recuperación, detección, segmentación y generación de imágenes. Asimismo, estará capacitado para proponer nuevos modelos y mover las fronteras de la práctica en esta área.* *La metodología de trabajo estará soportada por un 30% de cátedra para sentar los fundamentos en el área, un 40% de lectura y análisis de trabajos actuales (de CVPR, ICCV, NeuriIPS, IJCV, etc.), y finalmente un 30% de trabajos computacionales.*

Resultados de Aprendizaje	**Competencias a la que tributa (CE-CG)**
RA1: Entender los fundamentos de reconocimiento visual basado en deep learning.	CE1, CE2, CE9 CG1, CG2, CG3.
RA2: Entender y desarrollar modelos basados en deep learning para realizar búsqueda por similitud.	CE1, CE2, CE4, CE9 CG1, CG2, CG3, CG5, CG6
RA3: Entender y desarrollar modelos basados en deep learning para detección de objetos en imágenes.	CE1, CE2, CE4, CE9 CG1, CG2, CG3, CG5, CG6

RA4: Entender y desarrollar modelos basados en deep learning para segmentación de imágenes (segmentación semántica).	CE1, CE2, CE4 CE9 CG1, CG2, CG3, CG5, CG6
RA5: Entender y desarrollar modelos basados en deep learning para generación de imágenes.	CE1, CE2, CE4, CE9 CG1, CG2, CG3, CG5, CG6

Metodología Docente	Evaluación General
El profesor desarrolla la cátedra sobre los fundamentos en cada una de las temáticas definidas. Con los conocimientos adquiridos el estudiante estará en la capacidad de entender trabajos científicos en el área. Así, el estudiante deberá leer y realizar una presentación para un conjunto de trabajos actuales. Del mismo modo, para fortalecer los conocimientos teóricos, el estudiante deberá desarrollar trabajos experimentales sobre un conjunto de tareas relacionadas con los tópicos tratados.	• Control y examen: 40% El control y el examen tienen la misma ponderación. • Presentación de papers (4): 30% Todas las presentaciones tienen la misma ponderación. • Tareas computacionales (4): 30% Todas las tareas tienen la misma ponderación. El estudiante podrá ser eximido del examen siempre que se cumplan las siguientes condiciones : • Nota de control >= 5.5 • Promedio de presentaciones >= 5.5 • Promedio de tareas >= 5.5

Unidades Temáticas

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
1	RA 1	Fundamentos en Deep Learning en Visión Computacional	3
Contenidos		**Indicador de logro**	**Referencias a la Bibliografía**
• Redes Neuronales. • Redes Convolucionales. • Funciones de Activación. • Inicialización de Pesos y Regularización.		El estudiante es capaz de crear y evaluar modelos convolucionales para clasificación de imágenes.	1,2

• Métodos de Optimización basados en GD. • Redes Residuales y sus variantes.		

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
2	RA2	Búsqueda por Similitud	3
Contenidos		**Indicador de logro**	**Referencias a la Bibliografía**
• Redes Siamesas. • Redes Tripletas. • Extensión a Cuadrupletas • Generación de representaciones binarias (DeepHashing). • Redes basadas en grafos (GraphNet).		El estudiante es capaz de crear y evaluar modelos para búsqueda por similitud en imágenes.	1,3

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
3	RA3	Detección de objetos en imágenes	3
Contenidos		**Indicador de logro**	**Referencias a la Bibliografía**
• Faster RCNN y Anchors. • YOLO. • Retina Net. • TridentNet y convoluciones dilatadas (atrous convolution). • Modelos libres de anchors (FCOS). • Otros modelos de detección. • Modelos débilmente supervisados para detección de objetos.		El estudiante es capaz de crear y evaluar modelos para detección de objetos en imágenes.	4,5,6,7

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
4	RA 4	Segmentación de Imágenes	3
Contenidos		**Indicador de logro**	**Referencias a la Bibliografía**
• FCN. • U-Net. • Mask R-CNN. • DeepLab y sus variantes. • Modelos débilmente supervisados para segmentación semántica.		El estudiante es capaz de crear y evaluar modelos para segmentación de imágenes.	8,9,10

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
5	RA 5	Modelos Generativos	3
Contenidos		**Indicador de logro**	**Referencias a la Bibliografía**
• Variational AutoEncoders (VAE). • Generative Adversarial Networks (GANs). • Conditional GANs. • Cycle GAN. • GauGAN.		El estudiante es capaz de crear y evaluar modelos generativos.	1,11

Bibliografía General
Bibliografía obligatoria: [1] Ian Goodfellow and Yoshua Bengio and Aaron Courville. Deep Learning. 2016. (https://www.deeplearningbook.org/) [2] Christopher M. Bishop, “Pattern Recognition and Machine Learning”, Springer Verlag Gmbh, 2010, ISBN 978-0-387-31073-2. [3] *O. M. Parkhi, A. Vedaldi, A. Zisserman.* Deep Face Recognition. 2015. https://www.robots.ox.ac.uk/~vgg/publications/2015/Parkhi15/ [4] Faster-RCNN. https://arxiv.org/abs/1506.01497

[5] YOLO. https://arxiv.org/abs/1506.02640

[6] YOLOv3. https://arxiv.org/abs/1804.02767

[7] RetinaNet. https://arxiv.org/abs/1708.02002

[8] FCN. https://people.eecs.berkeley.edu/~jonlong/long_shelhamer_fcn.pdf

[9] UNet. https://arxiv.org/abs/1505.04597

[10] DeepLabv3+. https://arxiv.org/abs/1802.02611

[11] Pix2Pix. https://arxiv.org/abs/1611.07004

[12] Gilbert Strang. Linear Algebra and Learning from Data.

(https://math.mit.edu/~gs/learningfromdata/)

Bibliografía Complementaria:

Se revisarán trabajos publicados recientemente en importantes conferencias y revistas científicas, que incluyen a:

- CVPR
- ICCV
- NeurIPS

Vigencia desde:	2020
Elaborado por:	José M. Saavedra Rondo
Validado por:	Gonzalo Navarro y Sergio Ochoa
Revisado por:	