

## PROGRAMA DE CURSO

### VISIÓN COMPUTACIONAL

#### A. Antecedentes generales del curso:

Departamento	Ciencias de la Computación					
Nombre del curso	Visión Computacional	Código	CC5209	Créditos	6	
Nombre del curso en inglés	<i>Computer Vision</i>					
Horas semanales	Docencia	3	Auxiliares	1,5	Trabajo personal	5,5
Carácter del curso	Obligatorio		Electivo	x		
Requisitos	CC5205/CC4201					

#### B. Propósito del curso:

El propósito de este curso es proporcionar una comprensión integral de los fundamentos matemáticos y computacionales del procesamiento de imágenes y visión computacional, así como sus aplicaciones, destacando su impacto en áreas tecnológicas y científicas. Los estudiantes aplican algoritmos de detección de características, correspondencias geométricas y reconstrucción tridimensional para resolver problemas de percepción visual en escenas reales. Además, implementan soluciones basadas en aprendizaje profundo para tareas como clasificación, detección y segmentación de objetos.

En el curso se buscará hacer una evaluación crítica de diferentes enfoques de visión computacional, ayudando a seleccionar las técnicas más adecuadas según el problema y las restricciones computacionales. Finalmente, se desarrollan aplicaciones de software que integran componentes de visión computacional utilizando plataformas modernas, aplicando buenas prácticas de desarrollo y experimentación.

El curso tributa a las siguientes competencias específicas (CE) y genéricas (CG):

CE1: Analizar problemas computacionales, construir modelos, expresándolos en representaciones y lenguajes formales adecuados.

CE2: Analizar, diseñar y/o adoptar, algoritmos y estructuras de datos que cumplan con las garantías requeridas de correctitud y eficiencia.

CE5: Concebir, diseñar y construir soluciones de software, siguiendo un proceso sistemático y cuantificable, acorde a los fundamentos, eligiendo el paradigma y las técnicas más adecuadas.

CE6: Desarrollar software en una amplia variedad de plataformas y lenguajes

CG1: Comunicación académica y profesional:

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 una variedad de textos e informaciones sobre temas concretos y abstractos, comunicando experiencias y opiniones, adecuándose a diferentes contextos y a las características de la audiencia.



de programación.

**C. Resultados de aprendizaje:**

Competencias específicas	Resultados de aprendizaje
CE1, CE2	RA1: Explica los principales hitos, aplicaciones, fundamentos matemáticos y computacionales que sustentan las técnicas de procesamiento de imágenes y visión computacional, considerando el impacto de la visión computacional en diversas áreas tecnológicas y científicas.
CE1, CE2, CE5	RA2: Aplica algoritmos de detección de características, correspondencias geométricas y reconstrucción tridimensional para resolver problemas de percepción visual en escenas reales.
CE2, CE5, CE6	RA3: Implementa soluciones basadas en aprendizaje profundo diseñadas para tareas de visión computacional, tales como clasificación, detección y segmentación de objetos.
CE1, CE2, CE5	RA4: Evalúa críticamente el desempeño de distintos enfoques de visión computacional, seleccionando las técnicas más adecuadas en función del problema, los datos y las restricciones computacionales.
CE5, CE6	RA5: Desarrolla aplicaciones de software que integren componentes de visión computacional, considerando el uso de bibliotecas y plataformas modernas y aplicación de buenas prácticas de desarrollo y experimentación.
Competencias genéricas	Resultados de aprendizaje
CG1	RA6: Escribe reportes sobre el funcionamiento de algoritmos, el análisis de problemas en visión computacional, así como diseño de soluciones, considerando claridad, concisión, coherencia y uso de lenguaje técnico, propio del área de estudio, al momento de reportar sus resultados.
CG2	RA7: Lee de manera comprensiva, en inglés y español artículos académicos sobre métodos en visión computacional y su estado del arte, extrayendo información que aplica en la resolución de problemas de visión computacional.

**D. Unidades temáticas:**

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
1	RA1, RA2, RA4, RA6, RA7	Fundamentos de Visión Computacional	4 semanas
Contenidos		Indicador de logro	
1.1. Historia y aplicaciones de la visión computacional. 1.2. Fundamentos de imágenes digitales. 1.3. Procesamiento de imágenes (filtros, convoluciones, histogramas). 1.4. Detección de bordes y		El/la estudiante: 1. Explica los principales hitos y aplicaciones actuales de la visión computacional, identificando su impacto en distintas áreas tecnológicas y científicas. 2. Describe y analiza las características fundamentales de una imagen digital,	



<p>características (Canny, Harris, SIFT, SURF).</p> <p>1.5. Transformaciones geométricas y correspondencias.</p> <p>1.6. Calibración de cámaras y modelos de proyección.</p>	<p>incluyendo modelos de color, resolución espacial y profundidad de bits.</p> <p>3. Implementa técnicas básicas de procesamiento de imágenes como filtros, convoluciones y análisis de histogramas para mejorar o extraer información visual.</p> <p>4. Aplica algoritmos clásicos de detección de bordes y puntos de interés, evaluando su utilidad en distintos contextos de análisis de imágenes.</p> <p>5. Implementa operaciones geométricas sobre imágenes y aplica modelos de calibración de cámaras, justificando su uso en tareas de correspondencia y reconstrucción.</p> <p>6. Analiza problemas reales de visión computacional, sintetizando el estado del arte y escribe un reporte de análisis sobre el problema a resolver para el proyecto.</p>
<p><b>Bibliografía de la unidad</b></p>	<p>[1] Capítulo 1, 2, 4, 6. [4] Capítulo 3.</p>

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
2	RA1, RA2, RA4	Reconstrucción y Geometría	2 semanas
<b>Contenidos</b>		<b>Indicador de logro</b>	
<p>2.1. Estimación de movimiento y estructura desde movimiento.</p> <p>2.2. Epipolaridad, matrices fundamentales y esenciales.</p> <p>2.3. Visión estéreo, reconstrucción 3D básica.</p> <p>2.4. Introducción a SLAM.</p>		<p>El/la estudiante:</p> <ol style="list-style-type: none"> <li>Explica los principios teóricos de la geometría epipolar y su relación con la estimación de movimiento entre cámaras.</li> <li>Implementa algoritmos de reconstrucción 3D a partir de correspondencias estéreo, utilizando matrices fundamentales y esenciales.</li> <li>Analiza las ventajas y limitaciones de técnicas como estructura desde movimiento (SfM) y SLAM en escenarios de reconstrucción visual.</li> </ol>	
<b>Bibliografía de la unidad</b>		[2] Capítulo 9, 10, 11	



Número	RA al que tributa	Nombre de la unidad	Duración en semanas
3	RA1, RA3, RA4, RA5, RA6, RA7	Aprendizaje Profundo para Visión Computacional	3 semanas
Contenidos		Indicador de logro	
3.1. Redes convolucionales (CNN). 3.2. Arquitecturas clásicas: AlexNet, VGG, ResNet. 3.3. Transfer learning y fine-tuning. 3.4. Datasets: ImageNet, COCO, MNIST.		El/la estudiante: <ol style="list-style-type: none"> <li>Explica el funcionamiento de las redes convolucionales y el rol de sus componentes principales (convoluciones, pooling, activaciones, etc.) en la extracción de características visuales.</li> <li>Implementa y adapta arquitecturas clásicas de redes convolucionales (como AlexNet, VGG o ResNet) utilizando técnicas de transfer learning y fine-tuning en datasets estándar.</li> <li>Evalúa, a través de un análisis comparativo, el desempeño de distintos modelos CNN, sobre distintos datasets (MNIST, ImageNet, COCO), justificando la elección de arquitectura y estrategia de entrenamiento, según el problema a resolver.</li> <li>Analiza información, desde textos técnicos relacionados al uso de aprendizaje profundo para visión computacional.</li> <li>Escribe un reporte sobre el análisis de los datos, la arquitectura del modelo y la metodología de evaluación, respecto al problema a resolver para el proyecto del curso.</li> </ol>	
Bibliografía de la unidad		[3] Capítulo 9 [5] Capítulo 7, 8, 14	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
4	RA1, RA3, RA4, RA5, RA7	Tópicos Avanzados en Visión con Aprendizaje Profundo	6 semanas
Contenidos		Indicador de logro	
4.1. Detección y segmentación (YOLO, SSD, Mask R-CNN, UNet). 4.2. Detección de keypoints y estimación de poses. 4.3. Transformers visuales (ViT, DETR) 4.4. Modelos generativos (autoencoders, GAN's, modelos de difusión). 4.5. Aprendizaje de datos 3D.		El/la estudiante: <ol style="list-style-type: none"> <li>Explica los principios y diferencias clave entre arquitecturas avanzadas de detección y segmentación, incluyendo modelos basados en CNN y transformers visuales.</li> <li>Implementa y evalúa modelos de detección de objetos, segmentación semántica y estimación de poses utilizando frameworks actuales como YOLO y Mask R-CNN.</li> </ol>	



4.6. Neural Radiance Fields (NeRF).	<ol style="list-style-type: none"> <li>3. Diseña e implementa modelos generativos aplicados a datos visuales, comparando enfoques como autoencoders, GANs y modelos de difusión según el tipo de tarea.</li> <li>4. Implementa y evalúa una aplicación utilizando técnicas modernas como aprendizaje de datos 3D y Neural Radiance Fields (NeRF), evaluando su desempeño en tareas de reconstrucción o síntesis.</li> <li>5. Sintetiza información de textos técnicos, integrando dicha información a un reporte.</li> <li>6. Escribe el reporte final del proyecto, incluyendo las decisiones de diseño y los resultados de la evaluación de los algoritmos.</li> </ol>
Bibliografía de la unidad	<p>[3] Capítulo 20          [5] Capítulo 11,14, 20          [6] Capítulo 2, 3, 4</p>

### E. Estrategias de enseñanza-aprendizaje:

El curso considera las siguientes estrategias:

- **Clases expositivas:** se presentan en detalle los conceptos fundamentales de cada unidad. Los estudiantes resuelven pequeños problemas de implementación que permitan interiorizar los conceptos teóricos.
- **Uso de demos:** se utilizan herramientas tales como OpenCV o PyTorch, que facilitan el desarrollo y despliegue de aplicaciones para el área de estudio del curso.
- **Trabajo de proyecto:** se trabaja en un proyecto que resuelva un problema de complejidad intermedia. Se realizan 3 entregas que permitirán medir el grado de avance en el curso y el manejo de los conceptos en la práctica.

En las clases auxiliares los estudiantes trabajan con demos de los algoritmos vistos en clase y se resuelven problemas de análisis e implementación de variaciones de algoritmos de visión computacional.

### F. Estrategias de evaluación:

Al inicio de cada semestre, el académico o académica informará a los y las estudiantes sobre los tipos de evaluaciones, así como las ponderaciones correspondientes.

Tipo de evaluación	RA asociado a la evaluación	Ponderación
<ul style="list-style-type: none"> <li>■ Tareas: actividades de carácter práctico de temas vistos en clases.</li> </ul>	<p>SCC420e evalúan:</p> <ul style="list-style-type: none"> <li>■ Análisis sobre el uso de imágenes y su procesamiento, considerando la importancia de la visión computacional Se evalúa RA1.</li> <li>■ Aplicación de algoritmos de estimación de movimiento. Se evalúa RA2 y RA4.</li> </ul>	40%



	<ul style="list-style-type: none"> <li>■ Experimentación con Transfer Learning. Se evalúa RA3 y RA4.</li> <li>■ Aplicación de Transformers Visuales. Se evalúa RA1, RA3 y RA4.</li> <li>■ Experimentación con aprendizaje de datos 3D. Se evalúa RA3, RA4 y RA5.</li> </ul>	
■ Cuestionarios	Cinco cuestionarios a lo largo del semestre con preguntas de teoría y análisis. Se evalúa RA1, RA4 y RA7.	20%
■ Proyecto dividido en 2 – 3 iteraciones. Proyecto Grupal.	Actividades: <ul style="list-style-type: none"> <li>- Informes. Se evalúa RA6 y RA7.</li> <li>- Análisis del problema. Se evalúa RA1, RA4, RA6 y RA7.</li> <li>- Diseño de solución. Se evalúa RA2, RA3, y RA4.</li> <li>- Implementación de solución. Se evalúa RA4, RA5, RA6, y RA7.</li> </ul>	40%

**G. Recursos bibliográficos:**

<p><b>Bibliografía obligatoria:</b></p> <p>[1] Szeliski, Computer Vision: Algorithms and Applications (2da edición). Springer. 2022. Disponible online: <a href="https://szeliski.org/Book/">https://szeliski.org/Book/</a></p> <p>[2] Hartley y Zisserman, Multiple View Geometry in Computer Vision. Cambridge University Press. 2004.</p> <p>[3] Goodfellow, Bengio, Courville, Deep Learning. MIT Press. 2016. Disponible online: <a href="https://www.deeplearningbook.org/">https://www.deeplearningbook.org/</a></p> <p>[4] González y Woods, Digital Image Processing (4ta edición). Pearson. 2017.</p> <p>[5] Zhang, Lipton, Li, Smola, Dive into Deep Learning. Cambridge University Press. 2023.</p> <p>[6] Gao, Li, Deep Learning for 3D Point Clouds. Springer. 2025.</p>
--

**H. Datos generales sobre elaboración y vigencia del programa de curso:**

Vigencia desde:	Primavera 2025
Elaborado por:	Ivan Sipiran
Validado por:	Académico/a del área: Benjamín Bustos, Jefa Docente: Jocelyn Simmonds, CTD mes de mayo 2025
Revisado por:	Área de Gestión Curricular