

Carlos Henrique Caloete Pena

carloshenriquepena@gmail.com
chcp@cin.ufpe.br

+55 81 9 91685843
Brazilian/Portuguese

[linkedin.com/in/chcp/](https://www.linkedin.com/in/chcp/)
carlospena.com.br

SUMMARY

R&D Data Engineer/Scientist Specialist with over 8 years of experience and a Master's in Computer Science (UFPE), focused on delivering AI solutions that drive real business impact. I bring a solid foundation in **AI, computer vision, machine and deep learning**, with hands-on experience developing and deploying custom models at scale.

Beyond the technical side, I'm skilled at **leading** cross-functional teams and managing projects in **agile** environments aligning **AI** strategies with business goals and accelerating innovation. I work closely with stakeholders to turn complex challenges into production-ready solutions, combining deep technical knowledge with strong communication and leadership.

EDUCATION

Universidade Federal de Pernambuco (UFPE)

Master of Science in Computer Science (Segmentation of medical images - Advisor: Tsang Ing Ren)

Graduated in October 2022

GPA 4.0/4.0

- Publication on IEEE - IJCNN: [An Ensemble Learning Method for Segmentation Fusion](#)

Universidade Federal de Pernambuco (UFPE)

Bachelor of Science in Computer Engineering

Graduated in December 2019

GPA 8.67/10

ELS Language Centers - Vancouver, Canada

Intermediate level - General English Program

Graduated in June 2014

GPA 3.1/4.0

WORK EXPERIENCE

Ferreira Costa (Top 5 Retail/E-commerce in Brazil)

Recife, Pernambuco, Brazil

AI Architect (April 2025 - present),

July 2021 - Present

- **Lead and mentor** 6 cross-functional agile squads (~17 engineers, developers, data scientists, and analysts) in **R&D** of **AI-driven** solutions, including computer vision, data engineering, textual search and generative AI. Drive Scrum ceremonies, backlog refinement, and strategic alignment with Product Owners to ensure delivery is tightly connected to real business needs. While driving hiring, performance reviews, and technical career development.
- **Architected and led** a secure FastAPI/FastMCP-based platform exposing enterprise data (Oracle, GCP, internal APIs) to LLM agents and business teams via MCP-compatible interfaces, with authentication, rate limiting, and governance controls, enabling safe self-service data access without direct database exposure.
- **Architected** the development of a Bronze/Silver/Gold data platform using **SQLMesh** and **Dagster**, transforming unreliable SQL sources into trusted **BigQuery** datasets with automated testing, lineage, schema contracts, and orchestrated deployments, significantly improving data reliability and analytics readiness across teams.
- **Architected** the rollout of **semantic search** for long-tail queries, increasing CTR by +3.4 pp (+9.6% relative) using AWS SageMaker-hosted embeddings, with no material impact on response latency.

AI Data Scientist Specialist (jun 2024 -April 2025)

- **Lead** four agile squads in **R&D** of **AI-driven** solutions, including computer vision and generative AI. Drive Scrum ceremonies, backlog refinement, and strategic alignment with Product Owners to ensure delivery is tightly connected to real business needs.
- **Spearhead communication** between developers and stakeholders, transforming high-level objectives into technical tasks that solve customer pain points and support measurable outcomes.
- **Architected** a high-performance e-commerce **catalog** platform using **MongoDB** and **RabbitMQ**, enabling scalable ingestion and low-latency retrieval of product data.
- **Developed image processing pipeline** combining generative AI, deep learning, and traditional computer vision techniques. Utilized AWS SageMaker, integrating Python, OpenCV, FastAPI for automated product media enhancement and quality control.
- **Built and led** the development of chatbots and automated workflows using Generative AI with RAG, orchestrated via n8n.

Senior Data Scientist (Jul 2021 - May 2024)

- **Engineered** a 24/7 **Shipping System** capable of handling over 13K concurrent users during peak events like Black Friday. Ensured high availability with integrated observability tools and a tech stack including Python, Oracle, PostgreSQL, Redis, and AWS (ECR, EKS, S3).
- **Architected** and maintained high-performance **data structures (Materialized Views, scheduled Jobs, Redis caching layers)**, and **workflow orchestration** with Dagster
- **Oversaw** application **deployment pipelines** using GitLab CI/CD, AWS and Rancher, ensuring reliable version control and efficient CI/CD practices.
- **Built** a custom **product search engine** using Elasticsearch, significantly improving conversion rates by integrating insights from Google Analytics.
- **Honored** as **IT Professional of the Year** by FerreiraCosta/FCxLabs (2023)

NCR Tech Corporation (old name: OKI Brasil)

Recife, Pernambuco, Brazil

Graduate Research Fellow (Feb 2020 - June 2021), and Intern (Jun 2018 - Feb 2020).

June 2018 - June 2021

- **Researched and developed** computer vision and deep learning solutions.
- **Create and present weekly progress reports** to the client (audience: administrative and technical from NCR - São Paulo). Led discussions on emerging AI techniques, defining next steps based on new requirements, feature updates, and system limitations.

ACADEMIC EXPERIENCE

Centro de Informática (CIn) - Universidade Federal de Pernambuco (UFPE)

Recife, Pernambuco, Brazil

- *Scientific Research Scholarship Program* (object detection and image segmentation with deep learning)
- *Teaching Assistant* (Introduction to Programming and Digital Systems)

Jan 2017 - Jan 2019

Jan 2015 - Dec 2017

RobôCIn (UFPE Research Group)

Nov 2015 - Jul 2021

Co-founder of a research team with 80+ members developing autonomous robots.

- Led teams in robot soccer AI research, applying Deep Learning and Reinforcement Learning, [with publication on IEEE-LARS](#)
- Developed vision and control systems using OpenCV, QT, and ROS. For @Home (human-size) and IEEE VSSS (robot car)

SKILLS

PROGRAMMING LANGUAGES

Major: Python, SQL, PL/SQL, C/C++

SKILLS

Python, Computer Vision, Deep Learning, Machine Learning, Pytorch, AWS (EKS, SageMaker, S3), GCP (BigQuery), Git, GitLab CI/CD, Oracle, PostgreSQL, Elasticsearch, MongoDB, Redis, RabbitMQ, Docker, Agile, Leadership, Project Management, OpenCV, Generative AI, Dagster, Kubernetes, CI/CD, Data Engineering, Pandas, Polars, FastAPI, MCP

AWARDS

Honored as IT Professional of the Year by FerreiraCosta/FCxLabs (2023)

1st Place Team at Microsoft College Bot-ando Competition UFPE (2019)

Achieved top rankings in Latin American IEEE Very Small Size Soccer: 5th (2017), 4th (2020).

3rd Place Team at Microsoft College Code Competition UFPE (2018)

Top 10 project at National DragonBoard IOT embedded competition among 350 teams (2017)

Honorable Mention at Sigfox Embedded system - WND IoT Challenge (2017)

CERTIFICATIONS

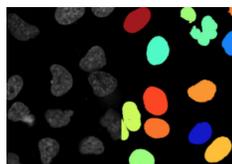
Data Engineering Specialization (DeepLearning.AI, Amazon Web Services, 2025); Data Storage and Queries (DeepLearning.AI, Amazon Web Services, 2025); Data Modeling, Transformation, and Serving (AWS, 2025); Source Systems, Data Ingestion, and Pipelines (DeepLearning.AI, 2025); Introduction to Data Engineering (AWS, 2025); Retrieval Augmented Generation (DeepLearning.AI 2025); LlamaIndex: creating a chatbot with the RAG technique (Alura 2024); Rust (Alura 2023); Kafka (Alura 2023); Agile Management Practices SC-AMP (Agile Institute Brazil 2022); Scrum (Alura 2022); Kanban Foundation KIKF (2020); Deep Learning with Pytorch (Udemy 2018); Deep Learning: GANs and Variational Autoencoders (Udemy 2017); 3D Printing with Fusion 360 (Udemy, 2017); Android (CITi, 2016); Logic: Language and Information (Melbourne, 2015).

LANGUAGES

Portuguese (Native)

English (Full professional proficiency)

PAPERS

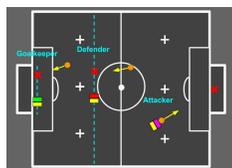


An Ensemble Learning Method for Segmentation Fusion

2022 IEEE - IJCNN, Padua-Italy, DOI: 10.1109/IJCNN55064.2022.9892717

A learning ensemble strategy that aggregates many independent candidate segmentations of the same image to produce a single consensus segmentation.

Keywords: Image Segmentation Fusion; Image Segmentation; Deep Neural Networks; Computer Vision.



An analysis of Reinforcement Learning applied to Coach task in IEEE Very Small Size Soccer

2020 IEEE - LARS, Natal-Brazil, DOI: 10.1109/IJCNN55064.2022.9892717

An end-to-end approach for the coaching task based on Reinforcement Learning, evaluated in the simulated environment of the IEEE Very Small Size Soccer (VSSS) competition.

Keywords: Reinforcement Learning; Neural Networks, Simulated Robots.

CONTACT

Email: carloshenriquepena@gmail.com | **Phone:** +55 81 9 91685843

LinkedIn: [linkedin.com/in/chcp](https://www.linkedin.com/in/chcp) | **GitHub:** github.com/carlospena00 | **Site:** carlospena.com.br

Location: Brazil | **Citizenship:** Brazilian/Portuguese