IntelliView

Your AI Interview Assistant

Meet our team!

Lauren Gallego

Junior Data Science & Engineering

🇪🇸

Alfonso Mayoral

Senior Telecommunications Engineer & Computer Science

🇪🇸

Pedro Tajia

Sophomore Data Science

🇲🇽

1

MEET THE PRODUCT

Overview & Key Features

1

Face Detection

A fine-tuned YOLO lightweight model is capable of drawing a bounding box over the user’s face

2

Emotion Detection

A second fine-tuned YOLO model is able to very accurately classify the user’s emotion from their facial expressions

3

Speech-to-text Transformation

Integrated Whisper model takes real-time audio and transcripts it into text form. This information is later used.

4

LLM Report Generation

Using the logged emotions and transcripted text, an LLM model generates a formatted report of the interview.

5

Real Time Analysis

Our application is able to do this in real time, with minimal latency that does not sacrifice any performance. This was one of the features that we prioritized.

Target audiences

Employers

Employers can use IntelliView to keep track of different candidates for different job positions using an independent automatic system.

Candidates

Candidates can use this tool as a practice tool to improve their interviewing skills, one of the main concerns of young potential candidates in today's market.

2

GOING INTO SPECIFICS

First Iteration

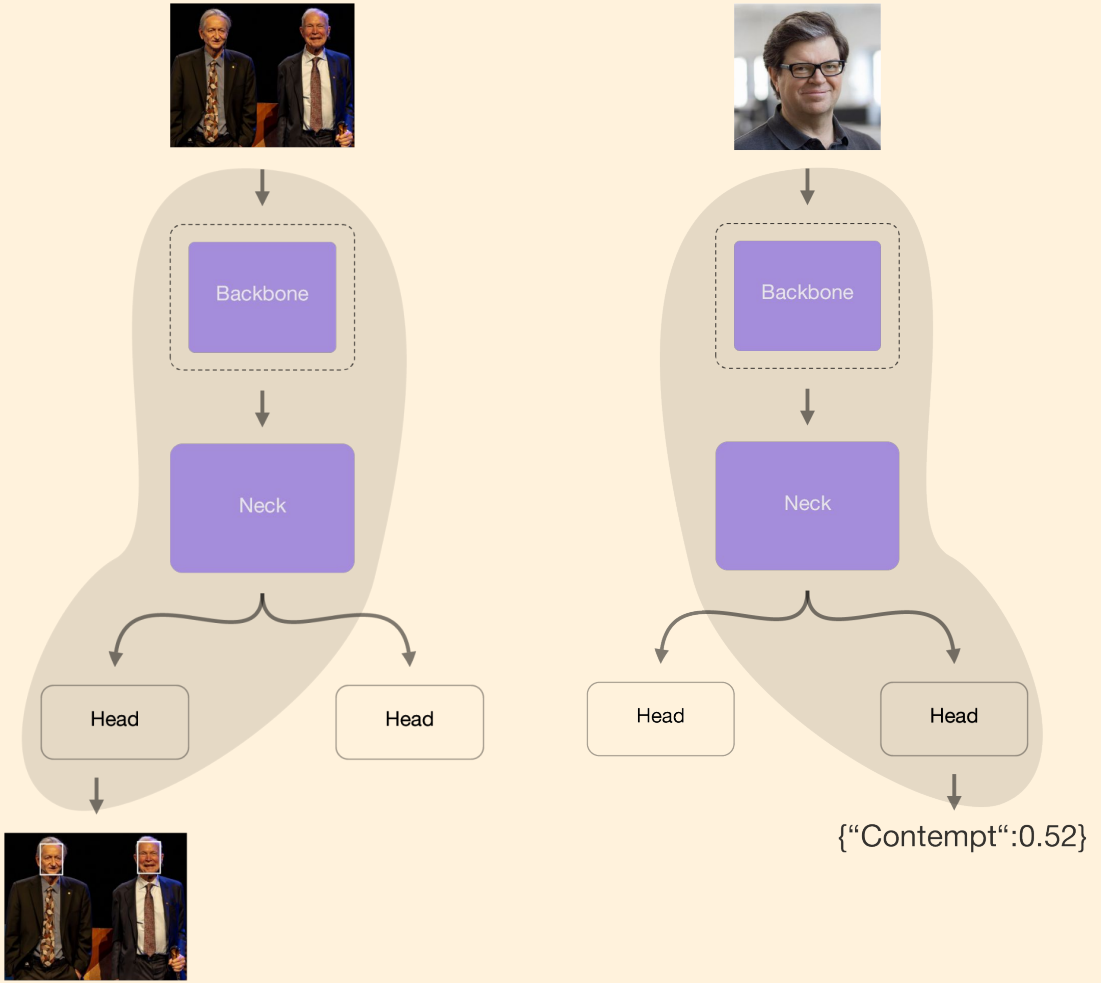
Have a system that detect a face and their emotion

Use a single yolo model, trained on two different datasets from **kaggle**.

- Facial Expression Image Data AFFECTNET YOLO Format
- Face-Detection-Dataset

Problem:

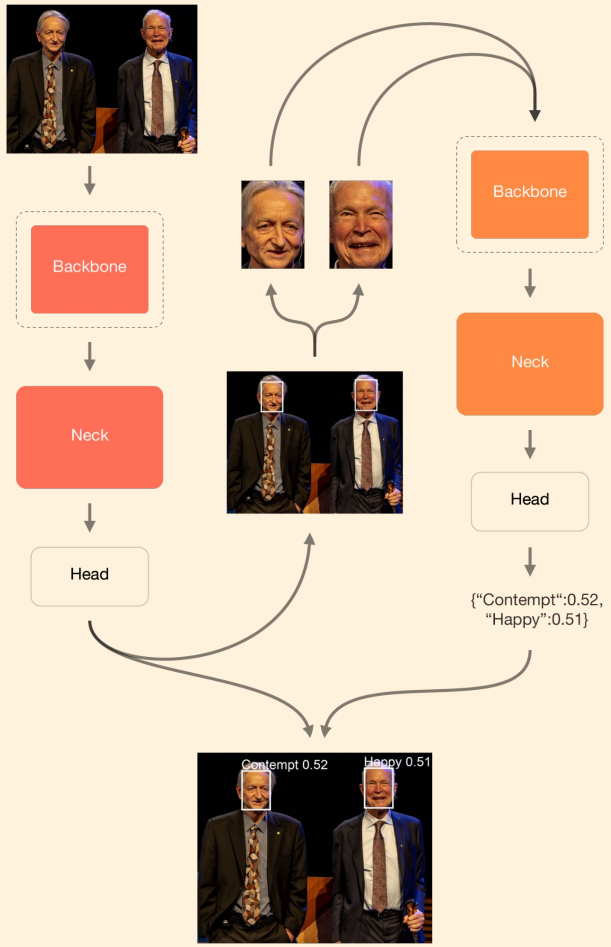
- Catastrophic forgetting
- Mismatch in the data



Final YOLO System

YOLO mode for **face** detection:

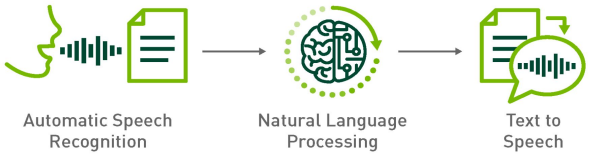
YOLO mode for **emotion** detection:



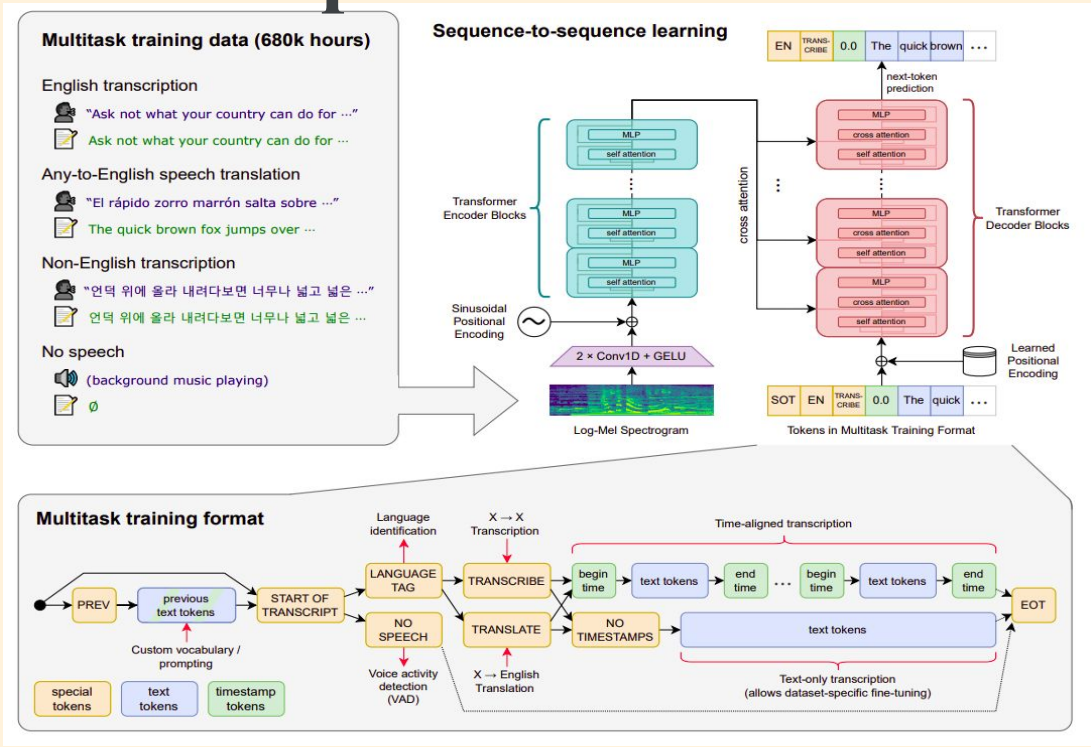
Real-time Speech to Text Capture

WHISPER

OpenAI's open-source Transformer trained on 680 000 hours of noisy, multilingual audio. This model listens, transcribes, and even translates speech in almost 100 languages as it happens. It outputs text tokens that carry 20 ms-resolution timestamps, so every word can be anchored to the exact video frame in which it is spoken, giving us an airtight bridge between sound, vision, and downstream analytics.

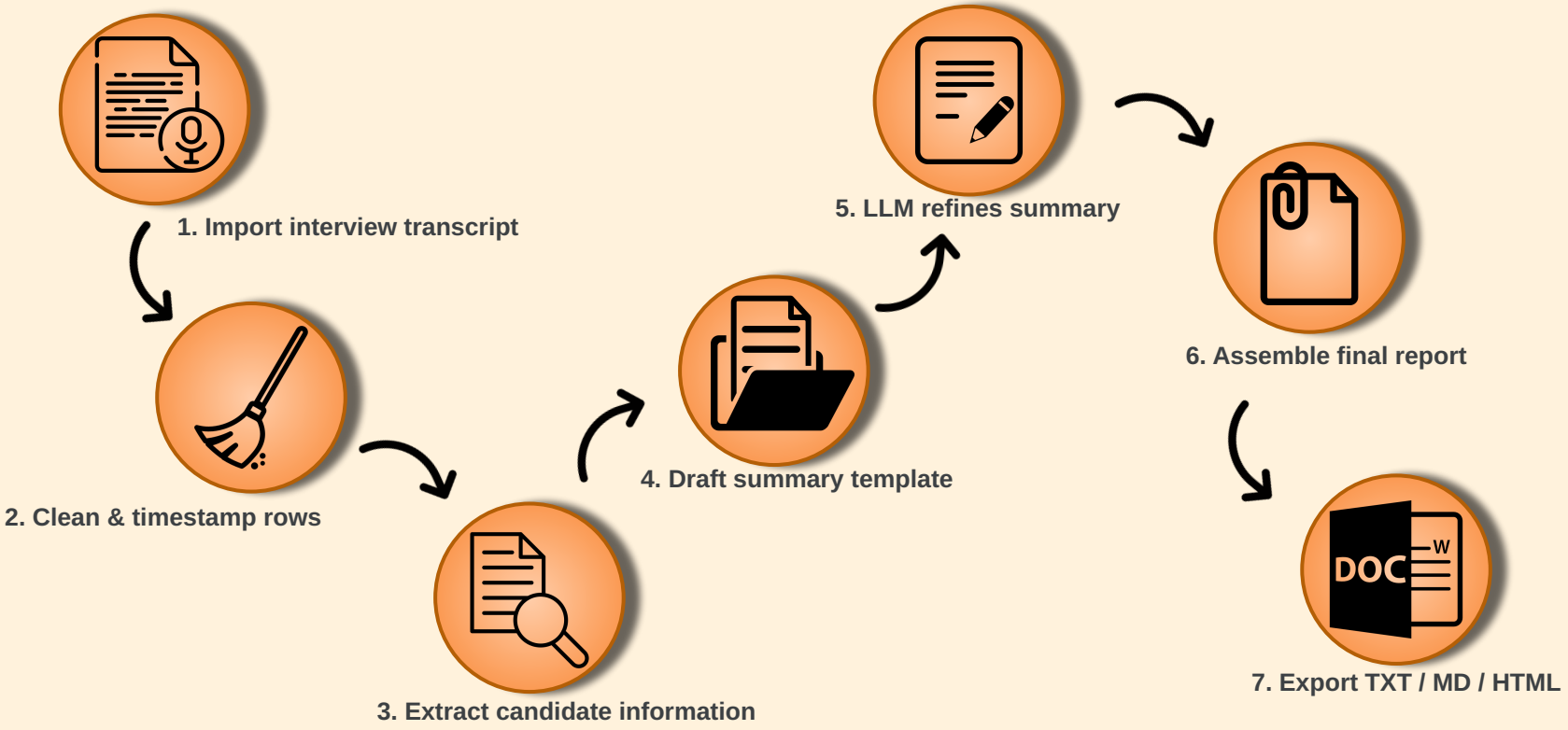


Model	Layers	Width	Heads	Parameters
Tiny	4	384	6	39M
Base	6	512	8	74M
Small	12	768	12	244M
Medium	24	1024	16	769M
Large	32	1280	20	1550M



[Whisper Official Paper]

LLM-Powered Interview-Report Agent



Evolution of agent development

1º Phase
Advanced models
(beginning of the project)

Models
- OpenAI GPT-4 Turbo 128k
- Mistral-7B-Instruct-v0.2
- **Llama 3-8B-Instruct**

Advantages
- Rich, human-like narrative that impressed recruiters.
- Could digest full 2-hour transcripts without trimming.

Disadvantages
- **10 minutes** average latency for free models due to computational cost
- $0.50 + per report with GPT-4.
- Very high computational and token cost

Next Steps:
- Cut latency & cost while keeping acceptable quality.
- Target on-prem inference to satisfy data-sovereignty rules.

2º Phase
Mid-size & Tiny Open Models + Heavy Prompting

Models
- **BART-Large-CNN SAMSum (Philschmid)**
- Llama-2-7b-chat-hf
- Microsoft Phi-3-mini-4k-Instruct
- Google Gemma-2B-IT

Advantages
- On-prem GPU deployment slashed cost ≈ 75 %.
- Tight prompts enforced consistent section headers.

Disadvantages
- Candidate-info extraction + external search added ~ 4mins, so overall latency stayed ≈ **7mins**.
- Language felt formulaic; occasional factual slips.

Next Steps:
- Embed fast, in-process keyword search functions to grab candidate facts and eliminate network lag; fold retrieved snippets straight into the prompt.

Final Phase
Search-Augmented Phi-1.5 (Current)

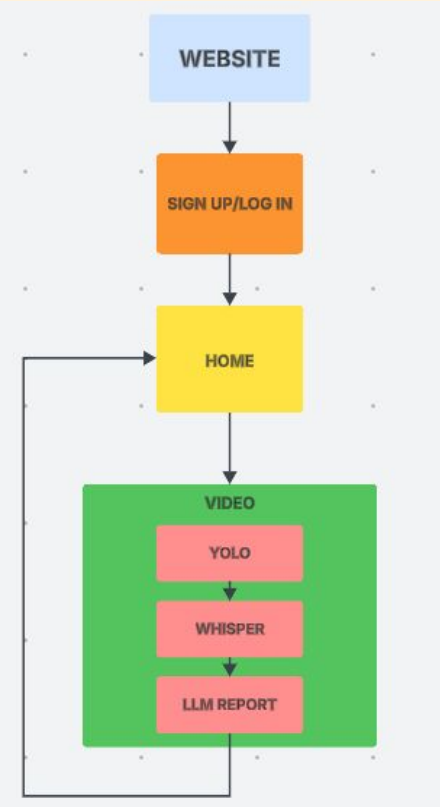
Models
- **Microsoft Phi-1.5** + in-process keyword search functions

Advantages
- **≈ 20 s** end-to-end thanks to regex-first extraction, strict 20 s summary timeout and 4-/8-bit quantisation.
- Runs on laptop CPU / 4GB GPU in harmony with the web app and the others models

Disadvantages
- Summary prompt must be extremely precise
- Narrative less rich and creative than advanced models
- 100–150-word limit and template fallback

Next Steps:
- Boost narrative quality without hurting speed.
- Fine-tuning or In-Context Learning
- Use advanced models and send notification via app when report is ready

Website architecture



Step one

1

Secure user integration using hashed passwords in order to create a personalized home page. Passwords have to satisfy a series of security conditions to be validated.

Step two

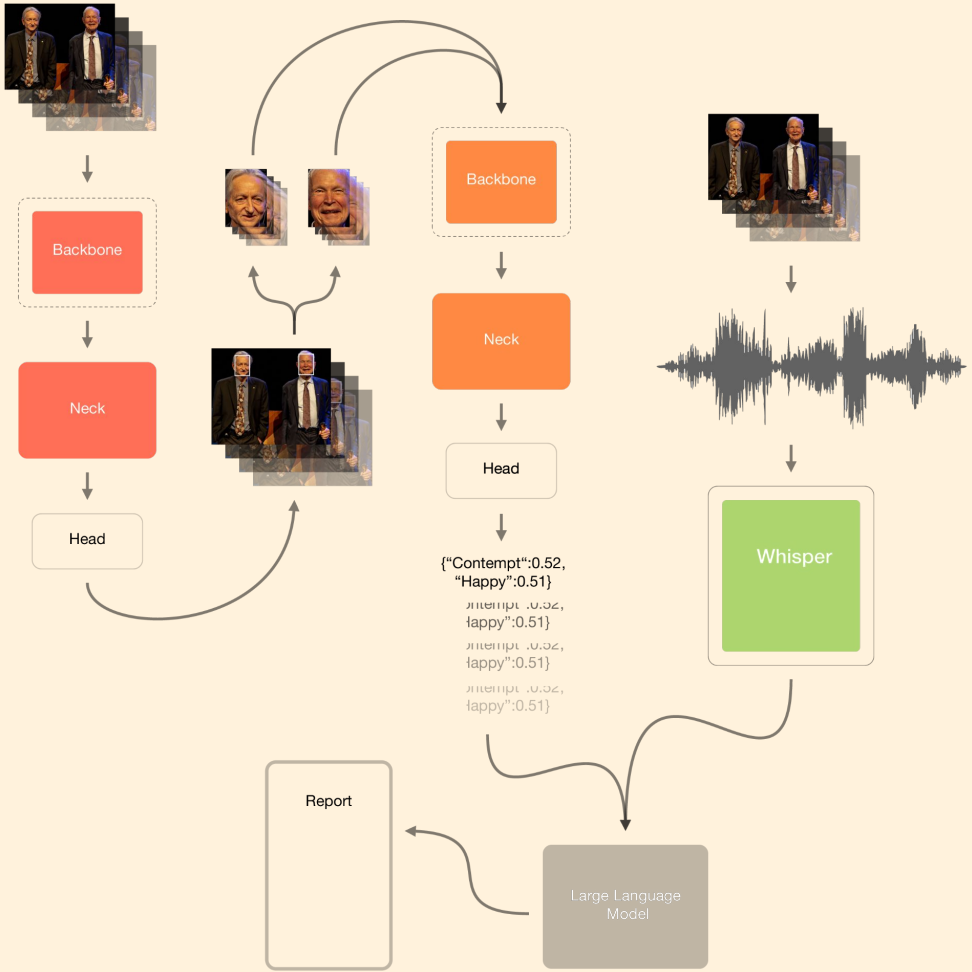
2

Access your own personal home page. In there you will be able to access, modify and delete your reports. These will be ordered by

Step three

3

Go into video, press ‘Start Recording’ and begin your interview!
Once you are done, press ‘Stop Recording’ and wait a couple of seconds for your automatic report to be generated


Backbone
Neck
Head
Backbone
Neck
Head
{“Contempt“:0.52,
“Happy”:0.51}
Whisper
Report
Large Language Model

3

IT’S TIME TO DEMO!

Comparison of Tested Models' Results (Llama 3 8B)

Candidate Information

- **Name**: Lauren Gallego
- **Professional Experience**: Data-Science Intern at Aitaca (Spain) – co-developed computer-vision models that predict jewellery ring size from hand images; collaborated in a cross-functional team and iterated rapidly on model performance and user-facing demo.
- **Academic Background**: B.Sc. in Data Science & Engineering – Universidad Carlos III de Madrid (2022 – May 2026, current). Exchange year in Computer Science at University of California, Riverside (2024-2025). Core focus on Machine Learning, Computer Vision and Artificial Intelligence; proficient in Python, SQL and R.
- **Strengths**: fast adaptability, solid ML/CV foundation, bilingual communication (Spanish/English), pro-active learning mindset, creative problem-solving, intercultural awareness.
- **Interview Duration**: 4 min 30 s

SUMMARY

The interview demonstrated that Lauren Gallego blends a solid academic foundation with directly relevant, hands-on experience and a proactive, adaptable mindset. He brings technical strengths in machine learning and computer vision that align with the role's requirements, while his bilingual communication skills and collaborative approach position him to contribute effectively within global, cross-functional teams.

- **Technical Experience & Project Leadership** – He outlines his internship at Aitaca, where he co-developed a computer-vision system that predicts ring sizes from hand images. His description highlights ownership of model experimentation and coordination with a cross-functional team, revealing emerging leadership skills.

- **Adaptability Example** – He recounts a 40 % scope change two weeks before launch: he helped restructure the project plan, pulled in extra resources, set up daily checkpoints, and still delivered on schedule—impressing the client and demonstrating composure under pressure.

- **Core Technical Competencies** – Proficient in Python (pandas, scikit-learn, PyTorch), SQL, and R; solid grounding in machine learning and computer vision; additional exposure to full-stack web development through a university project ("Intelligent Interview Assistant").

- **Creative Problem Solving** – He devised custom metrics to boost model accuracy and proposed UX flows that reduced user friction, illustrating innovative thinking beyond pure coding.

- **Multicultural Teamwork** – Having collaborated with peers in Spain and the U.S., he adapts communication styles to diverse audiences and values continuous feedback—skills vital for global teams.

- **Commitment to Continuous Learning** – Plans to pursue advanced deep-learning courses and an AWS Machine Learning certification; actively attends meetups and hackathons to stay current.

- **Communication & Empathy** – Communicates technical concepts clearly to non-technical stakeholders, listens actively, and structures responses logically—evidence of strong interpersonal skills.

- **Distributed-Team Experience** – Comfortable working across time zones using agile practices and collaborative tools (Jira, GitHub Projects), ensuring alignment and transparency.

Comparison of Tested Models' Results (BART-Large-CNN SAMSum)

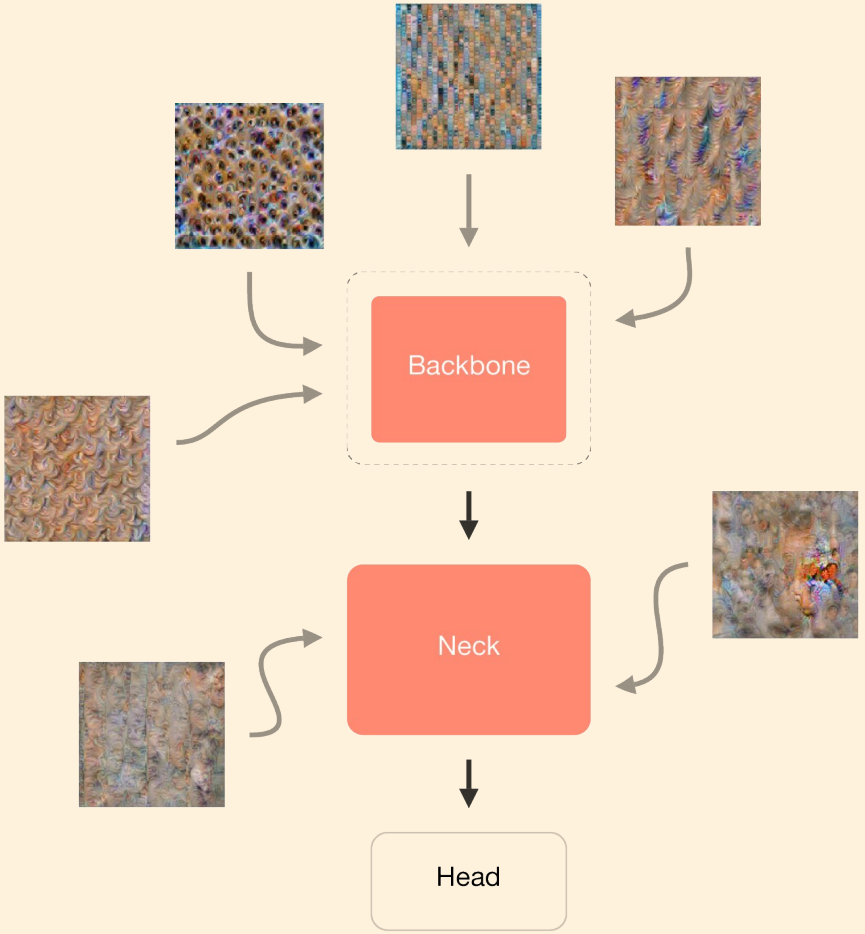
```
Name: Lauren Gallego
Date: 2025-05-27
Professional Experience: [{'company': 'Aitaca', 'job_title': 'Data Science Intern', 'duration': '3 months'}]
Academic Background: {'university': 'Universidad Carlos III de Madrid (exchange year at UC Riverside)', 'degree': 'B.Sc. in Data Science & Engineering (expected May 2026)'}
Skills: ['machine learning', 'computer vision', 'Python', 'SQL', 'R', 'adaptability', 'teamwork', 'communication', 'problem-solving']
Time: 4 minutes 30 seconds

SUMMARY:
Lauren Gallego is a final-year data-science student who brings hands-on experience from a computer-vision internship at Aitaca, where he helped build a model to predict
ring size from hand images. Proficient in Python, SQL and R, he focuses academically on machine learning and AI. During the internship he demonstrated adaptability by
reorganising a project after a late 40 % scope change, keeping the delivery on schedule. Bilingual in English and Spanish and comfortable in multicultural teams,
Lauren combines solid technical skills with clear communication and a proactive, problem-solving attitude—making him a strong prospect for an entry-level data or ML role.
```

AI Interpretability

- Understand the complexity of neural networks
- Building Trust and Confidence
- Mitigating Bias

Feature Visualization - Distill



Next Steps & Improvements

1. Run web application in a server
2. Use Whisper large models (1.55 B) for multilingual transcription
3. Use a more complex LLM that adapts to more open conversations
4. Create a more sophisticated web design
5. Train our own Emotion Classificator

5

ASK US ANYTHING!!

@aisc_madrid