

Proposal : AnomalibLVM / AnomalibGPT

Prerequisite (PR):

Openvino: [add opsets for BatchNorm in ONNX](#)

Anomalib: [Add MLFlow as a Logger in anomalib](#)

Anomalib: [Fix anomalib video, int image -> float](#)

Anomalib: [Add script to import custom folder.](#)

Hola! I'm Francisco, a Spaniard with a passion for AI and open-source innovation. My journey in Informatics Engineering began in Spain and culminated with a Master's in AI in Switzerland last February. During my academic endeavors, I've delved into the realms of Machine Learning, Deep Learning, and Computer Vision, improving my skills through practical applications and research.

Project Relevance and Personal Experience:

- My **Master's thesis** on leveraging image generation and computer vision to create a synthetic dataset for pest bug detection stands as a testament to my proficiency in these domains. This experience, aided by a course in Computer Vision and pattern recognition and another robotics course, has equipped me with practical skills in **Computer Vision**, and **Deep Learning** (using **PyTorch** and **PyTorch Lightning**), the necessary skills to tackle the challenges of integrating VLMs in Anomalib.
- My recent projects in **LLM** are <https://github.com/Bepitic/ActionArtisan>, in which I use the API of **OpenAI** and prompt engineering, to generate a description in a specific style. I also swap the API of this project to allow the utilization of **Hugging Face** models on: (<https://github.com/Bepitic/ActionArtisan-LocalLLM>), these 2 projects align perfectly with the project's requirement of integrating a VLM with the anomalib framework. This hands-on experience has provided me with deep insights into these models' capabilities and potential applications in various contexts.
- More Examples to showcase my proficiency in **Python** and familiarity with deep learning frameworks like **PyTorch Lightning** are gained both through academic projects and personal exploration, this can be seen in (<https://github.com/Bepitic/Data-Analytics>, and also I used PyTorch Lightning to train with **wandb** to do hyperparameter optimization on most of the homework-models in the Master), will be crucial in developing the model wrapper for Anomalib.
- Lastly, my expertise in C++ comes from my bachelor where I took multiple courses in advanced C++ focusing on low-level improvements, templates, and here: <https://github.com/Bepitic/LeetCode-Backup> where I save my solutions for leetcode problems in C++.

Data:

Name: Francisco Amoros Cubells

Uni: Università della svizzera Italiana (finished last feb)

Timezone: GTM +1 Madrid, Zurich, Paris.

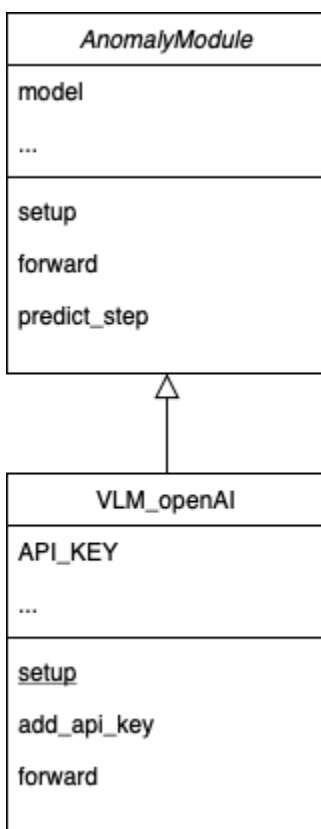
For the project I choose:

AnomalibVLM: Integrating Vision Language Models (VLMs) in Anomalib for zero- and few-shot anomaly detection.

Abstract

The project aims to integrate Visual Language Models (VLMs), using the API of OpenAI, into the anomalib framework to enhance anomaly detection using few-shot and zero-shot learning techniques.

Here is a figure outlining the design:



Starting with the creation of a model class 'LVM_openAI', that is descendant of 'AnomalyModule' in which I would add a couple of new parameters, `Api_key` (being a required one since openAI does not work without it), and another parameter 'prompt' (that is the prompt to send to the openAI, every dataset would require a custom prompt).

In this class I would overwrite the `setup` and `forward` functionality, `setup` to add the prompt and the `api_key` to the model, and `forward` to send the requests to the openAI API, and obtain the response (plus error catch for the API).

In order to make this more robust, adding tests is crucial, I would make a coverage test for the model 'LVM_openAI', that would

Test the override functions, and also the coupling with Pytorch lightning trainer.

Finally to Document and explain the functionality created, I will create a Notebook, with a toy example, explaining step by step how to set up th model and try it with this toy example, also the docstrings documenting the model and the functions will serve as the base for the API doc, and an explanation on the docs in the part of guides / models, explaining the reasoning and how to work with it with a python script for minimal setup.

After having this design implemented I will define a series of experiments in a notebook, to benchmark the performance against other models from Anomalib on multiple datasets (like MVTec AD and MVTec LOCO), among others.

Project Timeline

Before April 20:

- Deep dive into the Anomalib framework and understand its current anomaly detection techniques.

- Familiarize with OpenAI's ChatGPT API and its integration capabilities, research papers like ([GAD in VLM](#)), in order to understand how the landscape is at this moment.
- Design more in depth the functionality.

April 20 – May 26 (Pre-coding period):

- Develop prototypes integrating Anomalib with ChatGPT API.
- Try multiple prompts using ChatGPT UI for this task with multiple datasets.
- Engage with mentors and the community for advice on the design and the and refine the integration strategy.
- Prepare an experimental setup for testing the model on datasets like MVTec AD and MVTec LOCO.

May 26 – June 18 (Coding starts):

- Finalize and implement the model wrapper in Anomalib.
- Develop Python classes and functions for interacting with ChatGPT API.
- Begin initial testing and debugging.

June 18 – July 5:

- Intensive experiment testing with different datasets, prompts comparison with other models performances.
- Optimize the code based on initial test results.
- Start working on the optional UI component for model interaction.

July 6 – July 15 (Post Mid-Term):

- Refinement of the code for efficiency and accuracy.
- Detailed documentation of the developments and findings.
- Continual testing and bug fixes.

July 15 – July 25:

- Engage with the anomalib's developer community for feedback.
- Integrate any necessary modifications based on community input.
- Further refinement and testing.

July 25 – July 31:

- Finalize documentation.
- Prepare for the final evaluation and submission.

July 31 – August 19:

- Buffer period.

Optional Extensions (Post-GSOC):

- Replace ChatGPT API with a locally deployed open-source VLM such as LLaVa.
- Complete and refine the user interface for interactive model dialogues.

Skills and Experience Alignment

- Proficiency in Python and experience with deep learning frameworks like PyTorch, evident from previous project involvements.
- Solid background in machine learning and AI, with specific projects on image generation and text completion.
- Experience in software engineering, including full-stack development and cloud services integration.

Expected Outcomes

- A fully integrated VLM-based anomaly detection system within Anomalib.
- Comprehensive documentation and a set of experiments demonstrating the efficacy of the approach.
- An optional user interface enhancing the usability of the model.

General questions

- How/What do you know OpenVINO/anomalib?
I initially discovered OpenVINO and anomalib while exploring potential projects for Google Summer of Code. The philosophy behind these projects resonated with me, especially their focus on making AI more accessible by expanding GPU support to Intel's hardware lineup. This is in line with my passion for making AI and deep learning more inclusive and widespread.
- Have you already contributed to the OpenVINO/Anomalib project?
At the start of the doc.
- How could you apply it to your professional development?
Engaging with OpenVINO and anomalib presents a valuable opportunity for me at the onset of my professional journey as a Software/Machine Learning Engineer. Working on these projects would not only enhance my technical skills in AI and deep learning but also provide practical experience in applying these skills to real-world problems. This experience is pivotal in shaping my career trajectory, especially in terms of contributing to cutting-edge open-source projects and staying at the forefront of AI technology.
- Describe any other career development plan you have for the summer in addition to GSoC.
Apart from GSoC, I plan to continue expanding my knowledge base in AI and machine learning. Also I am currently searching for an open job opportunity in the field.
- Why should we pick you?
 - Strong Educational Background: My Master's in AI and Bachelor's in Informatics Engineering have equipped me with a solid foundation in AI, machine learning, and computer vision.
 - Relevant Project Experience: My experience with various LLM projects, including using the OpenAI API and substituting it with models from Hugging Face, directly aligns with the AnomalyGPT project's requirements.
 - Proven Skill Set: Proficiency in Python, deep learning frameworks (like PyTorch Lightning), and experience in cloud computing (AWS) ensure I can effectively contribute to and advance the project.
 - Commitment to Open-Source: My enthusiasm for open-source initiatives aligns with the spirit of GSoC, and I am keen on contributing to a project that has substantial real-world impact.
 - Innovative Approach: My past work, such as developing a synthetic dataset for pest bug detection using image generation and computer vision, demonstrates my ability to approach problems creatively and effectively.

Mentors: Samet Akcay, Dick Ameln, Ashwin Vaidya

Size of Project: 350 hours // Difficulty: Medium to Hard