

# The *COFE* Ecosystem

**GaNDLF**

Generally Nuanced  
Deep Learning  
Framework

**gandlf.org**

**ML**  
**Commons**

S. Pati, et al.,  
*Nature Communications Engineering*,  
2(23), 2023

**OpenVINO**

Model optimization for  
inference on low-  
resource environments

**openvino.ai**

**intel.**

A. Demidovskij, et al.;  
*ICCV Workshop*,  
783-787, 2019

**HF Hub**

Model deployment  
across multiple platforms  
& ecosystems

**hf.co**



S.M. Jain,  
*Introduction to  
Transformers for NLP*,  
51-67, Berkeley, 2022

**OpenFL**

Federated Learning  
Library

**openfl.io**

 **OPENFL**

P. Foley, et al.,  
*Phys Med Biol (ITCR  
Special Issue)*, 67(21),  
214001, 2022

**MedPerf**

Governance &  
Orchestration

**medperf.org**

**ML**  
**Commons**

A. Karargyris, et al.;  
*Nature Machine  
Intelligence*  
5:799-810, 2023

# Healthcare AI

Cyril Zakka, MD



**Hugging Face**

# About Me

**Founder & Program Director, AIM**

American University of Beirut Medical Center (2020)

**Medical Doctorate (MD)**

American University of Beirut Medical Center (2022)

**Postdoctoral Research Scholar**

Stanford Medicine (2024)

**Machine Learning Researcher, Health Lead**

HuggingFace



**Hugging Face**

# Overview

- What is HuggingFace?
- HuggingFace Health
- GaNDLF



**Hugging Face**

# Overview

- What is HuggingFace?
- HuggingFace Health
- GaNDLF



**Hugging Face**



# The AI community building the future.

The platform where the machine learning community collaborates on models, datasets, and applications.

**NEW** AI Tools are now available in HuggingChat

Tasks Libraries Datasets Languages Licenses Other

Filter Tasks by name

Multimodal

Text-to-Image  
Image-to-Text  
Text-to-Video  
Visual Question Answering  
Document Question Answering  
Graph Machine Learning

Computer Vision

Depth Estimation  
Image Classification  
Object Detection  
Image Segmentation  
Image-to-Image  
Unconditional Image Generation  
Video Classification  
Zero-Shot Image Classification

Natural Language Processing

Text Classification  
Token Classification  
Table Question Answering  
Question Answering  
Zero-Shot Classification  
Translation  
Summarization  
Conversational  
Text Generation  
Text2Text Generation  
Sentence Similarity

Audio

Text-to-Speech  
Automatic Speech Recognition  
Audio-to-Audio  
Audio Classification  
Voice Activity Detection

Tabular

Tabular Classification  
Tabular Regression

Reinforcement Learning

Reinforcement Learning  
Robotics

Models 469,541

Filter by name

meta-llama/Llama-2-70b

Text Generation • Updated 4 days ago • ↓ 25.2k • ❤ 64

stabilityai/stable-diffusion-xl-base-0.9

Updated 6 days ago • ↓ 2.01k • ❤ 393

openchat/openchat

Text Generation • Updated 2 days ago • ↓ 1.3k • ❤ 136

llyasviel/ControlNet-v1-1

Updated Apr 26 • ↓ 1.87k

cerspense/zeroscope\_v2\_XL

Updated 3 days ago • ↓ 2.66k • ❤ 334

meta-llama/Llama-2-13b

Text Generation • Updated 4 days ago • ↓ 328 • ❤ 64

tiiuae/falcon-40b-instruct

Text Generation • Updated 27 days ago • ↓ 288k • ❤ 899

WizardLM/WizardCoder-15B-V1.0

Text Generation • Updated 3 days ago • ↓ 12.5k • ❤ 332

CompVis/stable-diffusion-v1-4

Text-to-Image • Updated about 17 hours ago • ↓ 448k • ❤ 5.72k

stabilityai/stable-diffusion-2-1

Text-to-Image • Updated about 17 hours ago • ↓ 782k • ❤ 2.81k

Salesforce/xgen-7b-8k-inst

Text Generation • Updated 4 days ago • ↓ 6.18k • ↓ 57

# Models

<https://huggingface.co/models>



Hugging Face

Tasks   Libraries   Datasets   Languages   Licenses

Other

Filter Tasks by name

Multimodal

Image-Text-to-Text

Visual Question Answering

Document Question Answering

Video-Text-to-Text

Computer Vision

Depth Estimation

Image Classification

Object Detection

Image Segmentation

Text-to-Image

Image-to-Text

Image-to-Image

Image-to-Video

Unconditional Image Generation

Video Classification

Text-to-Video

Zero-Shot Image Classification

Mask Generation

Models 914,110   Filter by name

Full-text search   Sort: Trending

mattshumer/Reflection-Llama-3.1-70B  
Text Generation • Updated about 9 hours ago • 1.11k • 1k

black-forest-labs/FLUX.1-dev  
Text-to-Image • Updated 21 days ago • 756k • 3.89k

Qwen/Qwen2-VL-7B-Instruct  
Image-Text-to-Text • Updated 44 minutes ago • 121k • 421

Shakker-Labs/AWPortrait-FL  
Text-to-Image • Updated 1 day ago • 6.53k • 231

openbmb/MiniCPM3-4B  
Text Generation • Updated about 7 hours ago • 682 • 199

gpt-omni/mini-omni  
Text-to-Speech • Updated 2 days ago • 196

meta-llama/Meta-Llama-3.1-8B-Instruct  
Text Generation • Updated 17 days ago • 3.1M • 2.28k

# Datasets

<https://huggingface.co/datasets>



Hugging Face

Main Tasks Libraries Languages Licenses Other

Modalities

- 3D Audio Geospatial Image
- Tabular Text Time-series
- Video

Size (rows)

< 1K > 1T

Format

- json csv parquet
- imagefolder soundfolder
- webdataset text arrow

Datasets 205,800 Filter by name Full-text search Sort: Trending

- NousResearch/hermes-function-calling-v1  
Viewer • Updated 8 days ago • 11.6k • 1.28k • 161
- fka/awesome-chatgpt-prompts  
Viewer • Updated 3 days ago • 170 • 7.39k • 5.53k
- BAAI/Infinity-Instruct  
Viewer • Updated about 19 hours ago • 20.4M • 5.53k • 469
- opencsg/chinese-fineweb-edu  
Viewer • Updated 2 days ago • 30.7M • 55 • 39
- amphion/Emilia-Dataset  
Viewer • Updated about 7 hours ago • 52.9M • 2.34k • 51
- vidore/colpali\_train\_set  
Viewer • Updated 2 days ago • 119k • 980 • 25
- THUDM/LongWriter-6k  
Viewer • Updated 23 days ago • 6k • 634 • 140
- Salesforce/xlam-function-calling-60k  
Viewer • Updated Jul 19 • 60k • 1.71k • 350
- MarkrAI/KOpen-HQ-Hermes-2.5-60K  
Viewer • Updated 10 days ago • 60.1k • 306 • 44
- HuggingFaceM4/Docmatix  
Viewer • Updated 12 days ago • 2.55M • 889 • 183
- BAAI/Infinity-Preference  
Viewer • Updated 7 days ago • 59.4k • 142 • 24
- HuggingFaceFW/fineweb  
Viewer • Updated Jul 16 • 46B • 32.1k • 1.66k
- laion/relaion2B-en-research  
Viewer • Updated Jun 5 • 2.16B • 15 • 17
- Upstash/wikipedia-2024-06-bge-m3  
Viewer • Updated 10 days ago • 130M • 94 • 25



Discover amazing AI apps made by the community!

Create new Space

or [Learn more about Spaces](#)

<https://huggingface.co/spaces>

Search Spaces

Browse ZeroGPU Spaces ↗

Full-text search

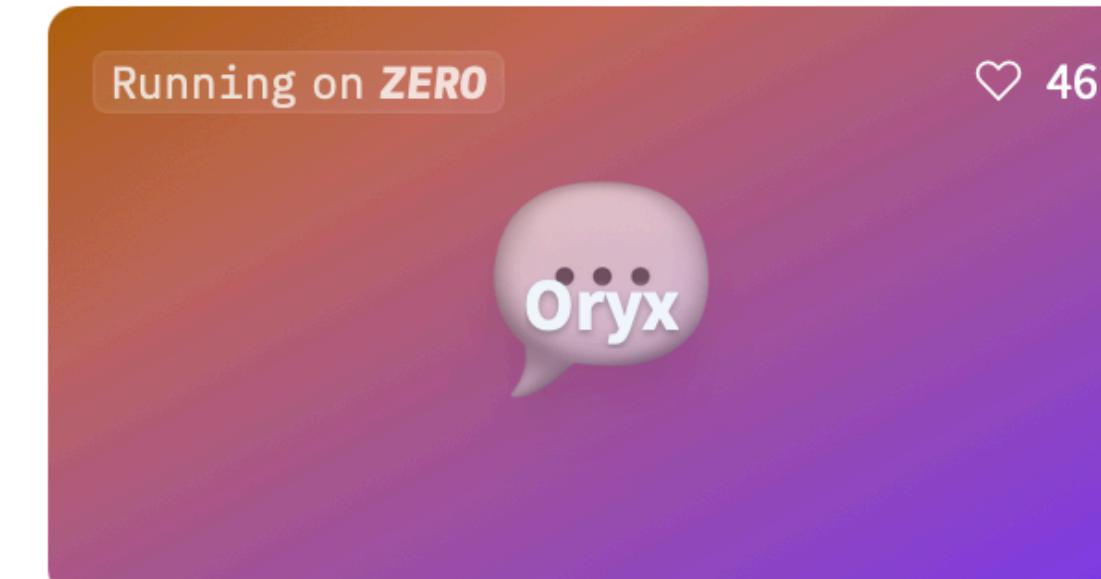
↑↓ Sort: Trending

### ☆ Spaces of the week 🔥



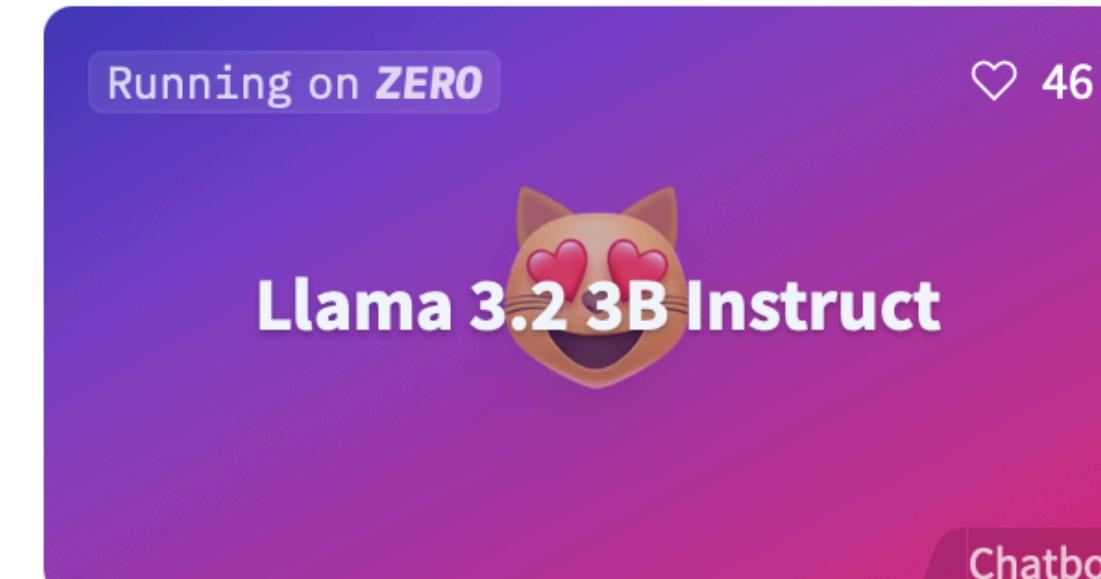
gabrielchua

about 4 hours ago



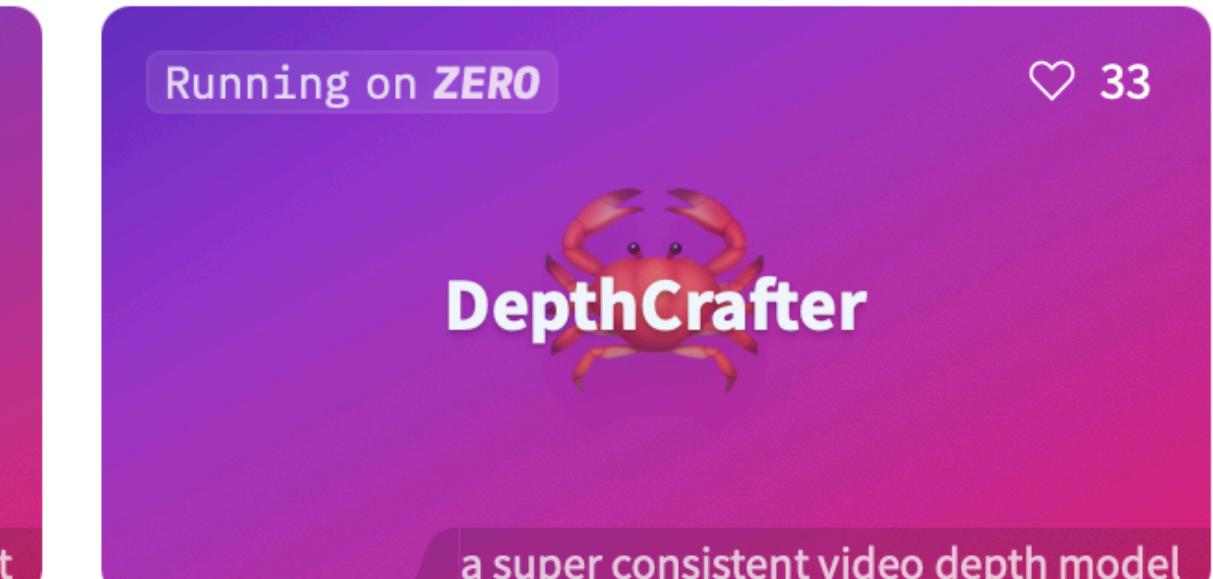
THUdyh

6 days ago



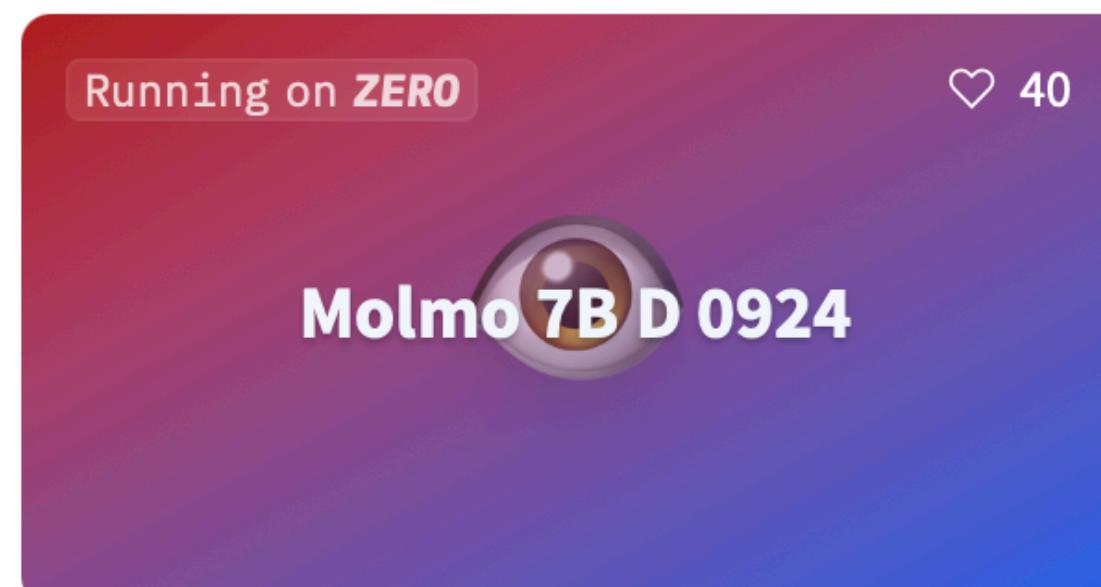
huggingface-projects

8 days ago



tencent

7 days ago



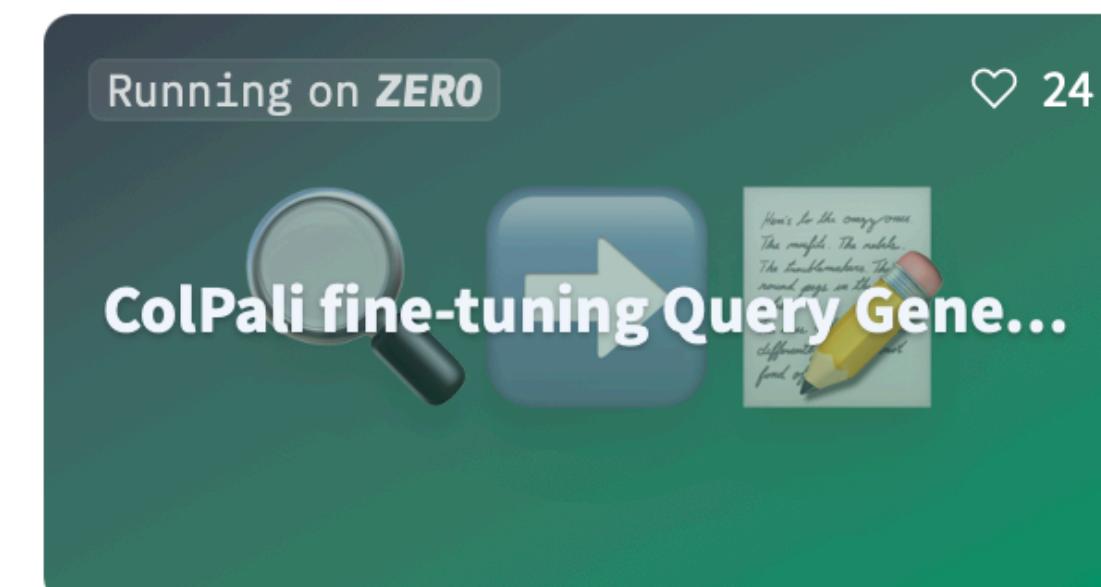
akhaliq

6 days ago



BAAI

1 day ago



davanstrien

5 days ago



alibaba-pai

1 day ago

# Transformers

<https://github.com/huggingface/transformers>



Hugging Face

README    Code of conduct    Apache-2.0 license    Security

 **Transformers**

build passing    license Apache-2.0    website online    release v4.45.1    Contributor Covenant v2.0 adopted    DOI 10.5281/zenodo.7391177

English | 简体中文 | 繁體中文 | 한국어 | Español | 日本語 | हिन्दी | Русский | Português | தெலுங்கு | Français | Deutsch | Tiếng Việt | العربية | اردو |

State-of-the-art Machine Learning for JAX, PyTorch and TensorFlow

Part of the

Zero-Shot Image Classification

OpenCLIP

English

clip

biology

medical

arxiv:2303.00915

License: mit

Model card

Files and versions

Community 25

:

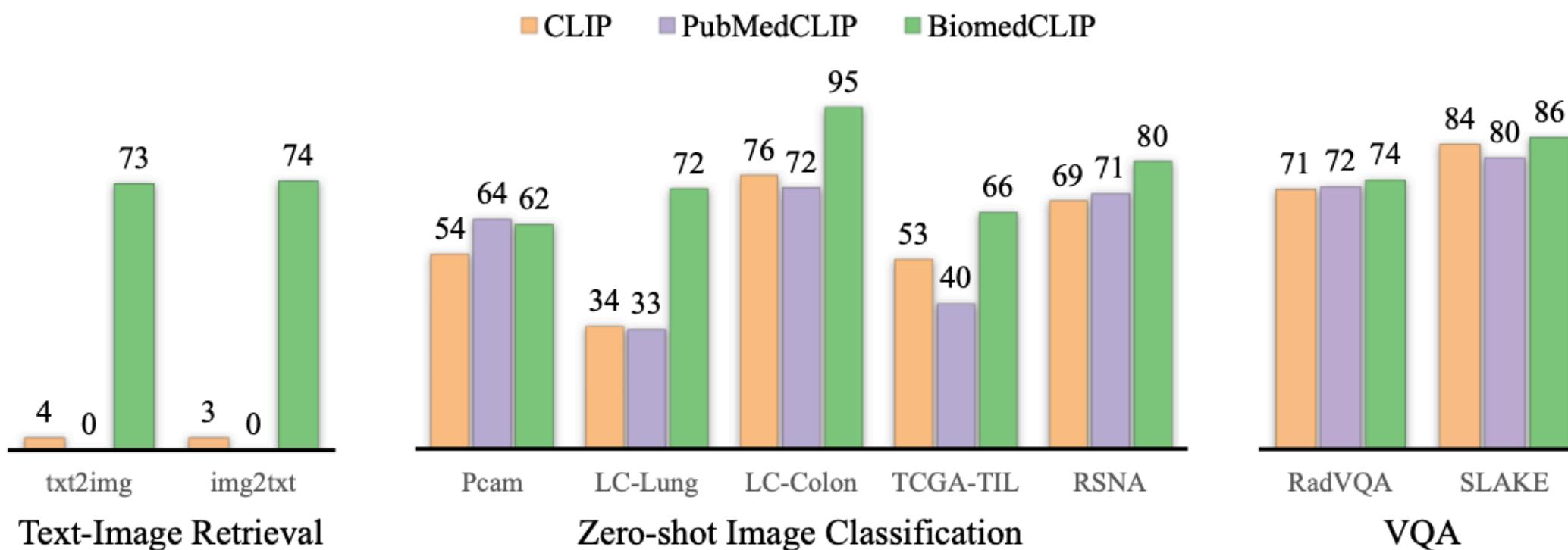
Use this model

## BiomedCLIP-PubMedBERT\_256-vit\_base\_patch16\_224

Edit model card

BiomedCLIP is a biomedical vision-language foundation model that is pretrained on PMC-15M, a dataset of 15 million figure-caption pairs extracted from biomedical research articles in PubMed Central, using contrastive learning. It uses PubMedBERT as the text encoder and Vision Transformer as the image encoder, with domain-specific adaptations. It can perform various vision-language processing (VLP) tasks such as cross-modal retrieval, image classification, and visual question answering.

BiomedCLIP establishes new state of the art in a wide range of standard datasets, and substantially outperforms prior VLP approaches:



Downloads last month  
28,442



## Inference Examples

Zero-Shot Image Classification

This model does not have enough activity to be deployed to Inference API (serverless) yet. Increase its social visibility and check back later, or deploy to [Inference Endpoints \(dedicated\)](#) instead.

## Spaces using microsoft/BiomedCLIP-PubMedBERT\_256-vit\_b...

- Baron-GG/LLAUS
- StefanDenner/MedicalVisualPromptEngineering
- zenitsu55/microsoft-BiomedCLIP-PubMedBERT\_256-vit\_base\_patch16\_224
- mitchmomo/microsoft-BiomedCLIP-PubMedBERT\_256-vit\_base\_patch16...
- charlestonX/microsoft-BiomedCLIP-PubMedBERT\_256-vit\_base\_patch16...
- Rajat456/microsoft-BiomedCLIP-PubMedBERT\_256-vit\_base\_patch16\_224
- pZacca/microsoft-BiomedCLIP-PubMedBERT\_256-vit\_base\_patch16\_224
- Dobator/microsoft-BiomedCLIP-PubMedBERT\_256-vit\_base\_patch16\_224

## Datasets: qiaojin/PubMedQA

Tasks: Question Answering Modalities: Text Formats: parquet Sub-tasks: multiple-choice-qa Languages: English Size: 100K - 1M ArXiv: arxiv:1909.06146  
Libraries: Datasets pandas Croissant + 1 License: mit

### Dataset card

Viewer

Files and versions

Community 3

#### Dataset Viewer

Auto-converted to Parquet API Embed Full Screen Viewer

Subset (3) pqa\_artificial · 211k rows Split (1) train · 211k rows

Search this dataset SQL Console

pubid	question	context	long_answer	final_decision
int32	string	sequence	string	string
25,429,730	Are group 2 innate lymphoid cells ( ILC2s )...	{ "contexts": [ "Chronic rhinosinusitis (CRS) is a..."	As ILC2s are elevated in patients with...	yes
25,433,161	Does vagus nerve contribute to the...	{ "contexts": [ "Phosphatidylethanolamine..."	Neuronal signals via the hepatic vagus nerv...	yes
25,445,714	Does psammaplin A induce Sirtuin 1-dependent...	{ "contexts": [ "Psammaplin A (PsA) is a natural..."	PsA significantly inhibited MCF-7/adr...	yes
25,431,941	Is methylation of the FGFR2 gene associated...	{ "contexts": [ "This study examined links between DNA..."	We identified a novel biologically plausible...	yes
25,432,519	Do tumor-infiltrating immune cell profiles and...	{ "contexts": [ "Tumor microenvironment immunity..."	Breast cancer immune cell subpopulation...	yes

Downloads last month

3,477

Use this dataset

Edit dataset card

Papers with Code

:

Homepage:

PubMedQA homepage

Repository:

PubMedQA repository

Paper:

PubMedQA: A Dataset for Biomedical Research Questio...

Leaderboard:

PubMedQA: Leaderboard

Size of downloaded dataset files:

300 MB

Size of the auto-converted Parquet files:

300 MB

Number of rows:

273,518

# Overview

- What is HuggingFace?
- **HuggingFace Health**
- GaNDLF



**Hugging Face**

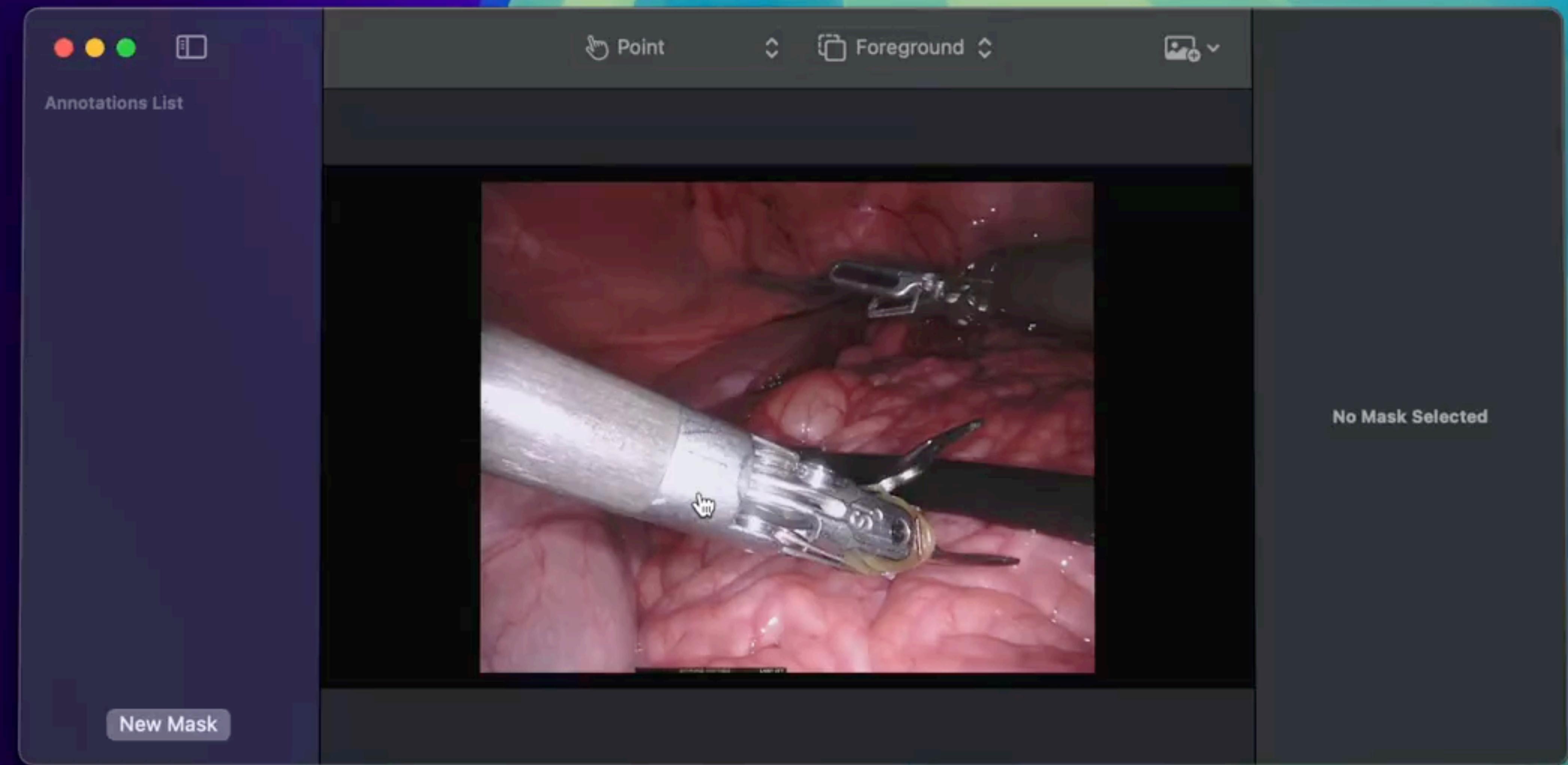
# Goals of HuggingFace Health

- Democratize medical and health AI research
- Develop tools to facilitate research and deployment
- Help train models via compute and expertise
- Train and OS medical models and artifacts of our own



**Hugging Face**





# Overview

- What is HuggingFace?
- HuggingFace Health
- **GaNDLF**



**Hugging Face**

# GaNDLF

```
def download_from_hub(  
    repo_id: str,  
    revision: Union[str, None] = None,  
    cache_dir: Union[str, None] = None,  
    local_dir: Union[str, None] = None,  
    force_download: bool = False,  
    token: Union[str, None] = None,  
):
```

```
def push_to_model_hub(  
    repo_id: str,  
    folder_path: str,  
    hf_template: str,  
    path_in_repo: Union[str, None] = None,  
    commit_message: Union[str, None] = None,  
    commit_description: Union[str, None] = None,  
    token: Union[str, None] = None,  
    revision: Union[str, None] = None,  
    allow_patterns: Union[List[str], str, None] = None,  
    ignore_patterns: Union[List[str], str, None] = None,  
    delete_patterns: Union[List[str], str, None] = None,  
):
```



**Hugging Face**



# Thank You

Email: [cyril.zakka@huggingface.co](mailto:cyril.zakka@huggingface.co)  
X: [@cyrilzakka](https://twitter.com/cyrilzakka)



**Hugging Face**