**Sprint 2:**

[**https://blog.roboflow.com/openai-clip/**](https://blog.roboflow.com/openai-clip/)

[**https://medium.com/@tastekinalperenn/yolox-main-idea-behind-latest-yolo-algorithm-5f8aa930c33c**](https://medium.com/@tastekinalperenn/yolox-main-idea-behind-latest-yolo-algorithm-5f8aa930c33c)

[**https://neptune.ai/blog/object-detection-with-yolo-hands-on-tutorial?utm\_source=facebook&utm\_medium=post-in-group&utm\_campaign=blog-object-detection-with-yolo-hands-on-tutorial&utm\_content=1854021431561644**](https://neptune.ai/blog/object-detection-with-yolo-hands-on-tutorial?utm_source=facebook&utm_medium=post-in-group&utm_campaign=blog-object-detection-with-yolo-hands-on-tutorial&utm_content=1854021431561644)

[**https://ai-scholar.tech/en/articles/self-supervised-learning/CoOp**](https://ai-scholar.tech/en/articles/self-supervised-learning/CoOp)

**Here the original at OpenAI with ground object as well as sat terrain example:** [**https://openai.com/blog/clip/**](https://openai.com/blog/clip/)

**Google AI Introduces ‘WIT’, A Wikipedia-Based Image Text Dataset For Visio-Linguistic Models**

**Quick Read:** [**https://www.marktechpost.com/2021/09/23/google-ai-introduces-wit-a-wikipedia-based-image-text-dataset-for-visio-linguistic-models/**](https://www.marktechpost.com/2021/09/23/google-ai-introduces-wit-a-wikipedia-based-image-text-dataset-for-visio-linguistic-models/)

**Github:** [**https://github.com/google-research-datasets/wit**](https://github.com/google-research-datasets/wit)

**Research:** [**https://dl.acm.org/doi/10.1145/3404835.3463257**](https://dl.acm.org/doi/10.1145/3404835.3463257)

**Also recommended this Chrome extension: Fatkun Batch Download Image, for large sets of image downloads, or automatic batch transformations, conversions etc**

[**https://towardsdatascience.com/using-image-segmentation-to-identify-rooftops-in-low-resolution-satellite-images-c791975d91cc**](https://towardsdatascience.com/using-image-segmentation-to-identify-rooftops-in-low-resolution-satellite-images-c791975d91cc)

[**https://medium.com/@tastekinalperenn/yolox-main-idea-behind-latest-yolo-algorithm-5f8aa930c33c**](https://medium.com/@tastekinalperenn/yolox-main-idea-behind-latest-yolo-algorithm-5f8aa930c33c)

[**https://laion.ai/laion-400-open-dataset/**](https://laion.ai/laion-400-open-dataset/)

[**https://blog.roboflow.com/openai-clip/**](https://blog.roboflow.com/openai-clip/)

[**https://9to5google.com/2021/09/29/googles-new-address-maker-app-will-let-communities-create-functional-addresses-quickly/**](https://9to5google.com/2021/09/29/googles-new-address-maker-app-will-let-communities-create-functional-addresses-quickly/)

[**https://www.androidauthority.com/google-lens-chrome-3032175/**](https://www.androidauthority.com/google-lens-chrome-3032175/)

[**https://youtu.be/LqAzpJthtIY**](https://youtu.be/LqAzpJthtIY)

[**https://developers.google.com/maps/documentation/api-picker?\_ga=2.221702170.386053195.1633094765-453183831.1632760629**](https://developers.google.com/maps/documentation/api-picker?_ga=2.221702170.386053195.1633094765-453183831.1632760629)

[**https://youtu.be/d1gNykSuauU**](https://youtu.be/d1gNykSuauU)

**According to this Google MUM featuring 75 languages already. Moreover, it has proven strong in a task on language variations, so it could potentially also work for the multilingual sub problem in the address normalizations**

[**https://lm.facebook.com/l.php?u=https%3A%2F%2Fventurebeat.com%2F2021%2F09%2F29%2Fhow-google-plans-to-improve-web-searches-with-multimodal-ai%2F&h=AT0L4tPJ3FlxxEBqFu3wdfYRf9p2tOc5CGAWT0Knjq4JWmkegtDXkArOluR8FIg3cdzmNwosUnm7c3Ui-oUl-NdTk9kaW86YKcLJhoE3QBYFWk-YRL0-i8VhkTaWX2Qc**](https://lm.facebook.com/l.php?u=https%3A%2F%2Fventurebeat.com%2F2021%2F09%2F29%2Fhow-google-plans-to-improve-web-searches-with-multimodal-ai%2F&h=AT0L4tPJ3FlxxEBqFu3wdfYRf9p2tOc5CGAWT0Knjq4JWmkegtDXkArOluR8FIg3cdzmNwosUnm7c3Ui-oUl-NdTk9kaW86YKcLJhoE3QBYFWk-YRL0-i8VhkTaWX2Qc)

[**https://stackoverflow.com/questions/54521080/aws-sagemaker-for-detecting-text-in-an-image**](https://stackoverflow.com/questions/54521080/aws-sagemaker-for-detecting-text-in-an-image)

[**https://www.marktechpost.com/2021/10/02/microsoft-ai-unveils-trocr-an-end-to-end-transformer-based-ocr-model-for-text-recognition-with-pre-trained-models/**](https://www.marktechpost.com/2021/10/02/microsoft-ai-unveils-trocr-an-end-to-end-transformer-based-ocr-model-for-text-recognition-with-pre-trained-models/)

**Baidu Research Introduces PP-LCNet: A Lightweight CPU Convolutional Neural Network With Better Accuracy And Performance**

**Quick Read:** [**https://www.marktechpost.com/2021/10/03/baidu-research-introduces-pp-lcnet-a-lightweight-cpu-convolutional-neural-network-with-better-accuracy-and-performance/**](https://www.marktechpost.com/2021/10/03/baidu-research-introduces-pp-lcnet-a-lightweight-cpu-convolutional-neural-network-with-better-accuracy-and-performance/)

**Paper:** [**https://arxiv.org/pdf/2109.15099.pdf**](https://arxiv.org/pdf/2109.15099.pdf)

**Github PaddleClas:** [**https://github.com/PaddlePaddle/PaddleClas**](https://github.com/PaddlePaddle/PaddleClas)

[**https://medium.com/augmented-startups/yolor-vs-yolox-battle-of-the-object-detection-prodigies-ae004a5ac8d2**](https://medium.com/augmented-startups/yolor-vs-yolox-battle-of-the-object-detection-prodigies-ae004a5ac8d2)

[**https://aws.amazon.com/about-aws/whats-new/2021/05/enhancements-to-amazon-rekognition-text-detection-support-for-more-words-higher-accuracy-lower-latency/**](https://aws.amazon.com/about-aws/whats-new/2021/05/enhancements-to-amazon-rekognition-text-detection-support-for-more-words-higher-accuracy-lower-latency/)

[**https://youtu.be/5QBM6d\_MrP4**](https://youtu.be/5QBM6d_MrP4)

**Multimodal Few-Shot Learning with Frozen Language Models (DeepMind)**

**"The idea this paper proposes is fairly simple: train a Language Model, freeze it such that its parameters remain fixed, and then train an image encoder to encode an image into a prompt for that language model to perform a specific task. I like to conceptualize it as “learning an image-conditional prompt (image through a NN) for the model to perform a task”.**

**Paper:** [**https://arxiv.org/abs/2106.13884**](https://arxiv.org/abs/2106.13884)

**Abstract:**

**"When trained at sufficient scale, auto-regressive language models exhibit the notable ability to learn a new language task after being prompted with just a few examples. Here, we present a simple, yet effective, approach for transferring this few-shot learning ability to a multimodal setting (vision and language). Using aligned image and caption data, we train a vision encoder to represent each image as a sequence of continuous embeddings, such that a pre-trained, frozen language model prompted with this prefix generates the appropriate caption. The resulting system is a multimodal few-shot learner, with the surprising ability to learn a variety of new tasks when conditioned on examples, represented as a sequence of multiple interleaved image and text embeddings. We demonstrate that it can rapidly learn words for new objects and novel visual categories, do visual question-answering with only a handful of examples, and make use of outside knowledge, by measuring a single model on a variety of established and new benchmarks."**

[**https://huggingface.co/transformers/model\_doc/clip.html#:~:text=CLIP%20is%20a%20multi%2Dmodal,to%20get%20the%20text%20features.**](https://huggingface.co/transformers/model_doc/clip.html#:~:text=CLIP%20is%20a%20multi%2Dmodal,to%20get%20the%20text%20features.)

[**https://www.google.com/intl/en-GB\_ALL/permissions/geoguidelines/**](https://www.google.com/intl/en-GB_ALL/permissions/geoguidelines/)

**Challenges for Google MUM:**

**MUM is Changing Search in Three Important Ways:**

[**https://www.google.ch/amp/s/www.searchenginejournal.com/google-search-redesign/421415/amp/**](https://www.google.ch/amp/s/www.searchenginejournal.com/google-search-redesign/421415/amp/)

**Things to know, Topic zoom, Visually browsable search results**

[**https://www.google.ch/amp/s/www.wired.com/story/soon-google-searches-combine-text-images/amp**](https://www.google.ch/amp/s/www.wired.com/story/soon-google-searches-combine-text-images/amp)

**SPRINT 1**

1. **Satellite image data with building detection on Open Buildings:**  
   Building footprints are useful for a range of important applications, from population estimation, urban planning, and humanitarian response, to environmental and climate science. This large-scale open dataset contains the outlines of buildings derived from high-resolution satellite imagery in order to support these types of uses. The current focus is on the continent of Africa with a coverage of about ⅔ of the whole continent and historic satellite image time series as well as confidence scores of the bounding area too. For more details: <https://sites.research.google/open-buildings>, which can be combined with the Google Earth Engine script to extract coordinates and directions.
2. **Alternative Global Building Footprints Extraction on the EOfactory platform:** (<https://eofactory.ai/>) at a faster rate  
     
   Identification and mapping of urban features such as buildings and roads are an important task for cartographers and #urbanplanners. Building footprint information is an essential component, and geospatial technologies help in creating this large mass of data inputs for designing and planning #smartcities.  
     
   With Machine Learning, extracting crucial information from imagery data is faster than the traditional approach, and saves cost, and has a high degree of accuracy. EOfactory is empowered with the advanced object detection algorithm which helps us to detect buildings, roads, and trees.  
     
   Trying the efficiency of EOfactory in supporting high-resolution aerial/ satellite images, and showcase with a use-case, the area used for EOfactory's Building footprint model over an Indian region of about 20 \*20 sq. km that just took 5 mins to process & give us quite accurate results of building footprints.  
     
   Do your own analysis by joining EOfactory, Join for free - <https://eofactory.ai/>  
     
   #EOfactory #Geospatial #GeoAI #geospatialintelligence #GeospatialDataScience #earthobservation #machinelearning #digitalimageprocessing #imageanalysis #remotesensing #earthscience #geoscience #geologist #geomatics #geologyscience
3. **Other global object and area detection via satellite FARM BOUNDARY MODEL:** (<https://eofactory.ai/>). Farm boundaries have been detected using edge detection training type. Edge detection allows users to observe the features of an image for a significant change in the gray level.  
     
   To assure the feasibility of the pre-trained models, We at EOfactory, have mastered the AI practices in Earth Observation and this is an example for farm boundaries of Haryana, with an approx. area of 44,212 sq. km that we have generated within 3 days:   
     
   #remotesensing #earthscience #earthobservation #geospatialdata #farmboundarydetection #ai #artificialintelligence #geospatial #geoscience #geologist #geology #earthsciences #imageryanalysis #satelliteimagery #satellitedata
4. <https://www.l3harrisgeospatial.com/Data-Imagery/Satellite-Imagery/High-Resolution/WorldView-3?gclid=Cj0KCQjwnJaKBhDgARIsAHmvz6dVI32Ood74C2p-azI9KjN7UMzEaPMdzX7K8zRp3wGV57oXV0Bu4YoaAtIHEALw_wcB>

This satelitte and image source seems to have the highest resolution at the moment and in about 29 spectra, including infrared, which would be the best for instance for letter or word detection etc

1. <https://medium.com/the-downlinq/you-only-look-twice-multi-scale-object-detection-in-satellite-imagery-with-convolutional-neural-34f72f659588>
2. Part I and II explain really well how to efficiently apply YOLO to satelite images for object detection via YOLT and they say they were able to localize infrastructure, such as airports accurately, and now also YOLTX more recently.

<https://towardsdatascience.com/announcing-yoltv4-improved-satellite-imagery-object-detection-f5091e913fad>: In its newer official version faster and more precise. We

1. Github.com/avanetten/yoltv4
2. Tutorial:

<https://towardsdatascience.com/announcing-yoltv4-improved-satellite-imagery-object-detection-f5091e913fad>

<https://m.youtube.com/watch?v=d1gNykSuauU&feature=share>