AI is mature, it's time to focus on clean datasets for the wind industry

Is a lack of high-quality training datasets holding the industry back from taking full advantage of the many benefits to be had from machine learning and Artificial Intelligence?

Why does AI matter in visual blade inspections?

The global wind fleet is set to increase to 3.7 TW by 2024, meaning more than 1.2 million wind turbines in operation. Traditionally, wind turbine blade inspections were performed from the ground, using high resolution cameras, telescopes and binoculars. Over the past five years, there's been a huge development in drone inspections, facilitating the capture of high quality images whilst minimizing turbine downtime.

Did you know that inspecting the entire global wind turbine fleet once a year means analyzing 400 million pictures every single year? This would result in hundreds of experts working full-time on this topic considering you need more than two minutes

per image. This is clearly not a viable option.

Artificial Intelligence plays a key role in reducing the overall inspection time, improving the efficiency and safety of wind farms on multiple levels. Centralized data platforms, paired with automated drone inspection flights and AI assisted defect detections can help to significantly reduce the time from image capture to detailed inspection report delivery, thus reducing maintenance and repair times.

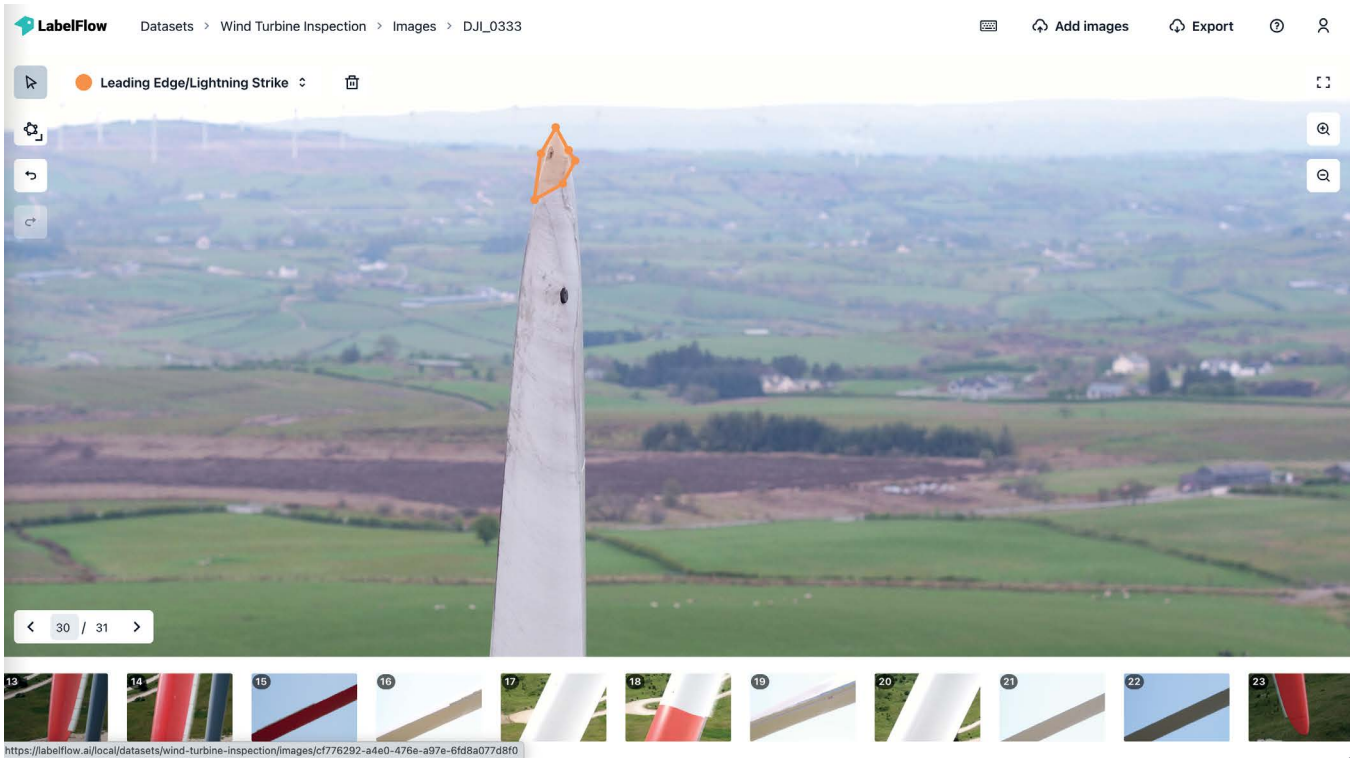
Specially trained AI for visual blade inspections paired with subject matter experts and drone pilots can usually reduce inspection times by up to 60 or 70%.

What's the status of AI for visual blade inspections in 2021?

There's been a shift recently in the machine learning community. Qualifying data and the pipelines to improve the performance of the AI has become more central than the machine learning algorithms themselves. Since the field has reached a certain level of maturity, the lack of high-quality training datasets are holding the industry back.

Improving the quality of training datasets is now the absolute key to improving computer vision capabilities, rather than fine-tuning the structure and parameters of the neural network, as there are now a number of benchmarks that can already validate the robustness of architectures.

In layman's terms, a training dataset is an initial set of data, which is used to help a program understand how to apply technologies such as neural networks to



learn and produce results. In the case of wind turbine inspection software with the objective of pre-labelling inspection images, the dataset will consist of expert-labeled images of wind turbine defects.

What makes a good dataset and how can you improve the quality?

Here are four aspects which will allow you to create great datasets:

High volume of data

No secret here, you need a high volume of labeled images to obtain great outputs with your machine learning model. Image augmentation can help you virtually increase this amount of data.

Balanced dataset

Size matters, but classes and homogeneity as well. If you have 1,500 labels for leading edge erosion and only 40 for lightning strikes, your model will probably fail at detecting this underrepresented class.

Training pipelines and toolings in place

A training pipeline needs to be set up to reach a virtuous circle for your AI. The process could be as follows: machine learning models create detections, or labels, on raw images. Human experts and labelers correct what's wrong or inaccurate, thus improving the training data quality. The machine learning models can then be trained once again on the improved quality dataset, and repeat.

Collaboration

Preparing a high-quality dataset often requires multiple stakeholders and a software platform to facilitate the work. This typically involves creating and modifying labels, merging labels into a larger dataset or even merging multiple datasets, amongst other actions. Data is dynamic, and should be continuously updated to maintain its integrity and relevance.

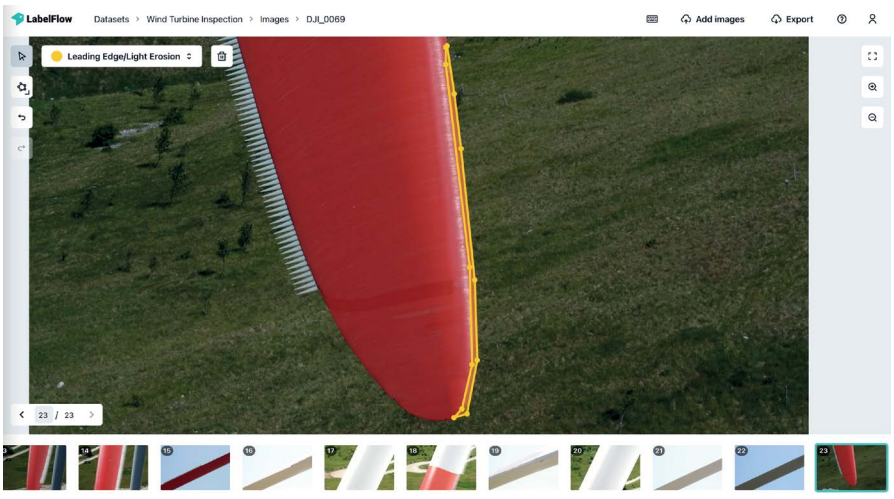
LabelFlow build the GitHub for visual data, with the mission to facilitate and accelerate the development of AI at scale, thanks to top-notch labeling tools and a dataset marketplace.

Launched in September 2021, designed around the smoothest user experience possible, we also opened our code on GitHub to empower the community and also benefit from wise user feedback. Over the last few years, LabelFlow technologies have supported dozens of projects, labeling more than 250,000 wind turbine images.

AI is just a tool, industry leaders have understood that the whole inspection value chain is important.

Over the past few years, multiple start-ups have entered the wind turbine inspection space, with varying focus points in order to capture different parts of the market.

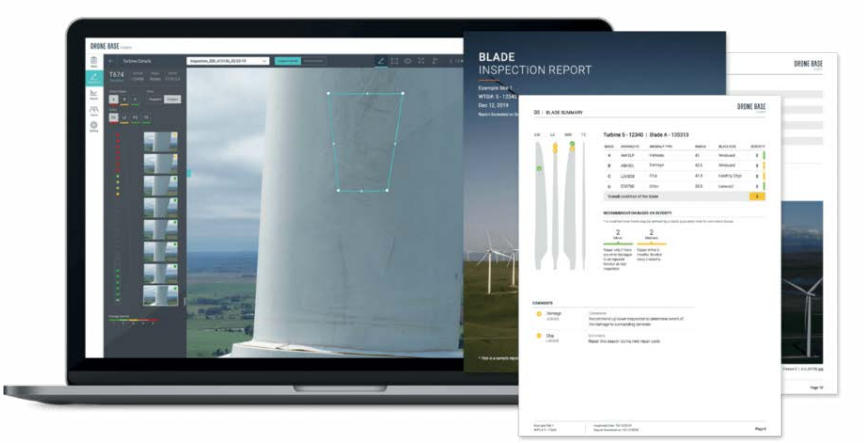
These industry leaders all successfully integrated AI, be it for automated flight planning software for drones or to automate turbine blade defect detections, but their biggest success is to have this integrated



into a larger workflow. This could include drone mission preparation, deployment, image sorting, analysis and curation, report generation i.e. PDF, CSV, etc., API integration, and all this with clean user interfaces, amongst other things.

Towards the end of 2021 we will publicly release some datasets to support the wind industry to move faster. If you are a solution vendor or a wind operator, feel free to get in touch with us if you want to be part of this open data project at contact@labelflow.ai

www.labelflow.ai



Here's our list of the top players in the smart wind turbine inspection ecosystem.

Skyspecs was founded in 2016, the North-American industry veteran has inspected over 33,000 turbines worldwide. Skyspecs combine the power of their proprietary drone technology, along with state-of-the-art Cloud-based asset management software to provide a best in class inspection service all over the world.

Dronebase's Wind Insights intelligent imaging platform delivers flexible, efficient power plant development and inspection solutions that provide immediate, actionable insights to improve risk mitigation and planning. With enterprise clients worldwide, DroneBase has successfully conducted over 150,000 commercial missions and is active in over 70 countries. We particularly appreciate their simple UI for streamlined image capture and analysis.

Scopito leverages their automation and machine learning expertise to derive all-important insights from your geospatial data, as well as concatenating historic data to enable powerful predictive maintenance. Scopito offers a streamlined SaaS asset management platform with a focus on simplicity, transparency and performance. Their easy onboarding process and no-frills approach is a welcome addition to the market.

SkyVisor is a relatively new player on the market, this young and dynamic French team offers AI powered data acquisition processing. Their solution for automated drone inspections has a particular focus on ease of deployment and productivity.

Cornis was founded in 2011, and having recently joined the SITES group, Cornis provides a complete set of tools for wind turbine inspections. Their portfolio includes ground-based camera technology with sub-millimeter precision, an automated drone inspection solution, as well as their Cornis Intrablade solution for detailed internal-blade inspections.

And the list is far from complete, with other players including **DroneDeploy**, **Cyberhawk**, **Aerodyne Measure**, **Sulzer & Schmidt**.