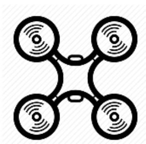
Capstone Project Proposal

Dan Chiriaev. *The "Super Drone Carrier"*

Product Sketch

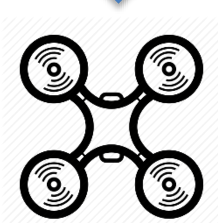


x10 drones

Each drone performs a function (e.g. inspection, delivery, maintenance, surveillance). Returns to carrier to recharge, download new job functions and schedule. No human interaction required.

Drone carrier does not necessarily fly. Can be stationary to reduce operational complexity and cost. It is a platform and large computer for drones to return and automate the job functions and scheduling of drones.

This is the main target of this project proposal.

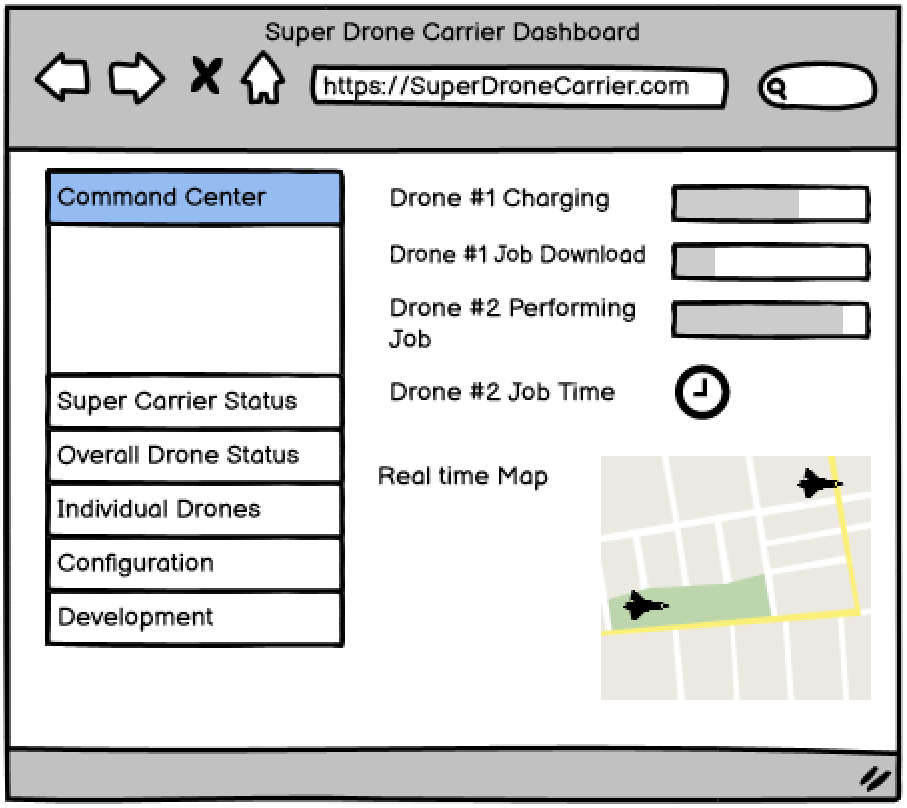


x10 drones

Drones proceed to next job. Drones can either be retrofitted or created from scratch to fit the super drone carrier product. The idea is that a drone can actually perform different job functions and it receives that training and instruction from docking at the carrier.

Product UI Sketch

(Intention is that this would be a desktop application, but could also be mobile for use in the field)



Business Goals

Project Overview and Goal What is the industry problem you are trying to solve? Why use ML/AI in solving this task? Be as specific as you can when describing how ML/AI can provide value. For example, if you're labeling images, how will this help the business?	Drones are a hot topic right now for various applications, but a common problem that any industry that uses drones is manual piloting. It requires several drone pilots and lacks automation. This project proposes a "super drone carrier" product can be developed, one that would allow automated flight scheduling, docking, charging, undocking, and automated job scheduling. ML/AI is proposed to solve this problem for image recognition (via drone cameras for anomaly detection), optimizing internal operations (automating various drone operations, e.g. docking, charging, job scheduling, flight scheduling), freeing up workers to be more creative (through reallocating job responsibilities from drone pilots from the actual flying to the monitoring of the drones), and enhancing current products (i.e. the drones themselves and the industry that benefits from using drones).
Business Case Why is this an important problem to solve? Make a case for building this product in terms of its impact on recurring revenue, market share, customer happiness and/or other drivers of business success.	This is an important problem to solve in terms of reducing maintenance of drones, number of hours spent manually flying drones, reducing operations cost, and increasing customer happiness by automating an increasing amount of workload.
Application of ML/AI What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve?	Image recognition (via drone cameras for anomaly detection), optimizing internal operations (automating various drone operations, e.g. docking, charging, job scheduling, flight scheduling), freeing up workers to be more creative (through reallocating job responsibilities from drone pilots from the actual flying to the monitoring of the drones), and enhancing current products (i.e. the drones themselves and the industry that benefits from using drones).

Success Metrics

Success Metrics What business metrics will you apply to determine the success of your product? Good metrics are clearly defined and easily measurable. Specify how you will establish a baseline value to provide a point of comparison.	-Flight time hours (either increase or decrease depending on the use-cases at hand) -Maintenance hours (goal is to reduce) -Human working hours (goal is to reduce) -Operating costs (goal is to reduce) **V2.0 Edit:** **Establishing a baseline:** **1) Flight time hours are measured compared to the pre-deployment environment, for all job functions, for all human pilots.** **2) Maintenance hours are measured based on current number of hours worked prior to deployment.** **3) Human working hours are similarly measured as maintenance and flight time hours** 4) **Operating costs are measured on a pre-deployment baseline. The baseline may be averaged out over a year's timeframe.**

Data

Data Acquisition Where will you source your data from? What is the cost to acquire these data? Are there any personally identifying information (PII) or data sensitivity issues you will need to overcome? Will data become available on an ongoing basis, or will you acquire a large batch of data that will need to be refreshed?	The data is to be intended to be acquired on an ongoing basis through various sensors built in the drones themselves, such as figuring out how much space should be between drones when they land on the super carrier. In practice, the data above could be acquired from drones we purchase and use ourselves in building the product or the data could be found through open source means. The product would have to feature significant computational power and sensors which would generate a lot of data on its own after it were created.

	Edit V2.0: **For personally identifying information:** **The tricky part is that drones need to recognize what human faces look like in order to scrub the faces in a production environment. Consultation with data privacy and computer vision experts would be needed to ensure that PII is not accidentally transmitted or misused.**
Data Source Consider the size and source of your data; what biases are built into the data and how might the data be improved?	If we find open source data on drones, we need to consider what job function the drone was intended for. Because a drone that is used to perform oil and gas line inspections will generate data that isn't suitable for delivery purposes. Essentially, we want drones to be able to dock at the super carrier and be able to intake new jobs (similar, not exact, and not too different) when needed. This means we need multiple data sets of training so the same drone could receive them from the super carrier. **Edit V2.0:** **There can definitely be a few types of bias in the project that need to be recognized.** **The problem with a big project like this that requires hardware and software integration is there is a real possibility for measurement bias between the training data used and the production data. This means the rollout team needs to have quality control embedded in the product to monitor satisfactory progress towards the success metrics.** **Another foreseeable bias is association bias. If the super drone carrier is used for delivery purposes, a dataset that contains fast food delivery only to working class neighborhoods would problematically reinforce a cultural stereotype. As far as the model would be concerned, middle and upper class neighborhoods don't order fast food, even though that's not true in real life.**
Choice of Data Labels What labels did you decide to add to your data? And why did you decide on these labels	Data labels have to be there for specific job functions that the drones will receive from the carrier. That means each specific function has its own data sets with its respective labels. E.g. we want the super drone carrier

versus any other option?	to be able to tell drones to do inspection and delivery jobs. That means the training data has to have data labels specific to those two functions. **Edit V2.0:** **Let's use the inspection example.** **For inspection, data labels may be:** **-Starting and Ending XYZ coordinates** **-Correct state of object (e.g. pipeline, boring and long)** **-Anomaly (break down into subcategories)** **a) Accident (e.g. leak, explosion)** **b) Trespassing (e.g. someone unauthorized walking on pipeline)**

Model

Model Building How will you resource building the model that you need? Will you outsource model training and/or hosting to an external platform, or will you build the model using an in-house team, and why?	The challenge is to find external model training that would essentially teach a product to switch between different models (e.g. carrier programs drone to do 2 different or more type of jobs, not just 1 job) In-house would certainly be more expensive and slower to do. Such a team would have to build a series of models that would be programmed to trigger depending on the job task at hand. The series of models would then programmed into the brain of the super drone carrier that assigns job functions to individual drones and the carrier receives orders from a control room. **Edit V2.0: An external tool like Google AutoML can be used to build small experimental training models and then combined with in-house development.**
Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?	Precision and recall would be metrics that need to be evaluated based on the job function of the drone. If a drone does multiple functions, the model metrics for all job functions have to be measured and monitored. The super carrier performance needs to be measured in how fast it assigns jobs, how well it trains new drones to do new jobs, how much energy it costs to run this carrier, etc. Because this is a pretty break through idea, the first

	version would not be as polished and perhaps not perform as intended. It’s the next few versions down the line that would be expected to perform at a higher level. **Edit V2.0:** **Foreseeing a costly product, I believe the standard baselines for such a product need to be very high, like 95-98% because it will serve as important hub for customers where lots of important work gets done.** **If a drone does only a single job function, like delivery, the precision level has to be high, e.g. 98% +. The reason for that is because something like delivery is a high-volume business, and each % point could mean misplaced customer packages. But, we also need to be practical and realistic.** **For a multi-job function: The paradigm I see is the collective performance rate needs to be passable, e.g. 70% precision, while new versions of the product are iterated and training/production data is improved for specific job functions.**

Minimum Viable Product (MVP)

Design What does your minimum viable product look like? Include sketches of your product.	MVP needs to be a platform of any shape (TBD, could be circular, rectangular) that allows multiple drones to dock, automatically charge, receive new programs to perform new jobs, undock, and remain connected to the super carrier. New compatible drones may need to be developed or existing drones need to be retrofitted to work with the super carrier. <SKETCHES>
Use Cases What persona are you designing for? Can you describe the major epic-level use cases your	Use cases: -Jobs that use drones to collect data and use that data to improve product -Roles that use indoor or outdoor inspections -Visual mapping, e,g, vegetation, forestry, water,

product addresses? How will users access this product?	construction sites -Delivery (commercial and medical) -Maintenance Persona: Companies that use multiple drones at the same time as part of their operations, whether for maintenance, mapping, delivery, inspection, and/or data collection. Access to product: The super drone carrier is accessed through a combination of wireless hardware and software programs.
Roll-out How will this be adopted? What does the go-to-market plan look like?	The roll out has to target industries and companies that see heavy drone usage, for these are the users that would benefit the most from automation and AI/ML. **Edit V2.0:** **1) Hire the team.** **There needs to be a product owner, a product designer, a UI/UX designer, a QA, a ML engineer, a data annotator, a data scientist, a 2-3 software engineers, and 2-3 hardware engineers. The hardware engineers and product designer need to come first to build the physical product. Hires should be done within 1 month.** **2) Generate datasets** **These can be obtained from open source means, generated in-house, or obtained through purchase. Labeling would have to occur either through an external data annotation company and/or internally through a data annotator.** **3) Build the physical product** **We expect this process to take about 1-2 months due to the complexity of the product. This may take longer and require more hardware engineers because the product is a breakthrough idea.** **4) Build the software and integrate with physical product** **Programming and connecting software with**

	hardware should take around 1 month. **5) Build and train models** **First iteration should take around 2-3 weeks** **6) Test model, iterate until precision meets specification** **I would expect 2 weeks spent on getting this phase to production level.** **7) Launch** **Begin deployment to first customers. Engage sales force. Monitor current version performance. Iterate new version. Provide customer support and patches.** **The total time from scratch to launch should be in the range of 4-6 months.** **8) Iteratively improve product** **Add new datasets and retrain models on a monthly basis.**

Post-MVP-Deployment

Designing for Longevity How might you improve your product in the long-term? How might real-world data be different from the training data? How will your product learn from new data? How might you employ A/B testing to improve your product?	The product would have to feature significant computational power and sensors which would generate a lot of data on its own after it were created. That data would need to be collected and used to improve various operations and efficiencies. A/B testing can be used in the super drone carrier uploading 1 job versus 2 jobs to drones. E.g. 50% of drones may be asked to do only 1 type of job (say, maintenance) and the other half are asked to do maintenance and delivery type of jobs. Data from

	both groups would be compared and analyzed for insights and business action.
Monitor Bias How do you plan to monitor or mitigate unwanted bias in your model?	An obvious source of bias comes from the super drone carrier not having enough training/production data for various job functions to upload to drones. Example: an oil and gas company was using the super drone carrier to upload maintenance jobs for drones to complete. When the same carrier is asked to assign off-shore or on-shore delivery from oil rigs, the drones may not have enough training and production data to do the job well.