

DEEP INDUSTRIAL ESPIONAGE

Samuel Albanie, James Thewlis, Sébastien Ehrhardt & João F. Henriques

Dept. of Deep Desperation,
UK (EU at the time of submission)

ABSTRACT

The theory of deep learning is now considered largely solved, and is well understood by researchers and influencers alike. To maintain our relevance, we therefore seek to apply our skills to under-explored, lucrative applications of this technology. To this end, we propose and *Deep Industrial Espionage*, an efficient end-to-end framework for industrial information propagation and productisation. Specifically, given a single image of a product or service, we aim to reverse-engineer, rebrand and distribute a copycat of the product at a profitable price-point to consumers in an emerging market—all within in a single forward pass of a Neural Network. Differently from prior work in machine perception which has been restricted to classifying, detecting and reasoning about object instances, our method offers *tangible business value* in a wide range of corporate settings. Our approach draws heavily on a promising recent arxiv paper until its original authors' names can no longer be read (we use felt tip pen). We then rephrase the anonymised paper, add the word “novel” to the title, and submit it a prestigious, closed-access espionage journal who assure us that *someday*, we will be entitled to some fraction of their extortionate readership fees.

1 INTRODUCTION

In the early 18th Century, French Jesuit priest François Xavier d’Entrecoulles radically reshaped the geographical distribution of manufacturing knowledge. Exploiting his diplomatic charm and privileged status, he gained access to the intricate processes used for porcelain manufacture in the Chinese city of Jingdezhen, sending these findings back to Europe (over the course of several decades) in response to its insatiable demand for porcelain dishes (Giaimo, 2014). This anecdote is typical of corporate information theft: it is an arduous process that requires social engineering and expert knowledge, limiting its applicability to a privileged minority of well-educated scoundrels.

Towards reducing this exclusivity, the objective of this paper is to democratize industrial espionage by proposing a practical, fully-automated approach to the theft of ideas, products and services. Our method builds on a rich history of *analysis by synthesis* research that seeks to determine the physical process responsible for generating an image. However, in contrast to prior work that sought only to determine the parameters of such a process, we propose to instantiate them with a *just-in-time*, minimally tax-compliant manufacturing process. Our work points the way to a career rebirth for those like-minded members of the research community seeking to maintain their *raison d'être* in the wake of recent fully convolutional progress.

Concretely, we make the following four contributions: (1) We propose and develop *Deep Industrial Espionage* (henceforth referred to by its cognomen, *Espionage*) an end-to-end framework which enables industrial information propagation and hence advances the *Convolutional Industrial Complex*; (2) We introduce an efficient implementation of this framework through a novel application of differentiable manufacturing and sunshine computing; (3) We attain qualitatively state-of-the-art product designs from several standard corporations; (4) We sidestep ethical concerns by failing to contextualise the ramifications of automatic espionage for job losses in the criminal corporate underworld.

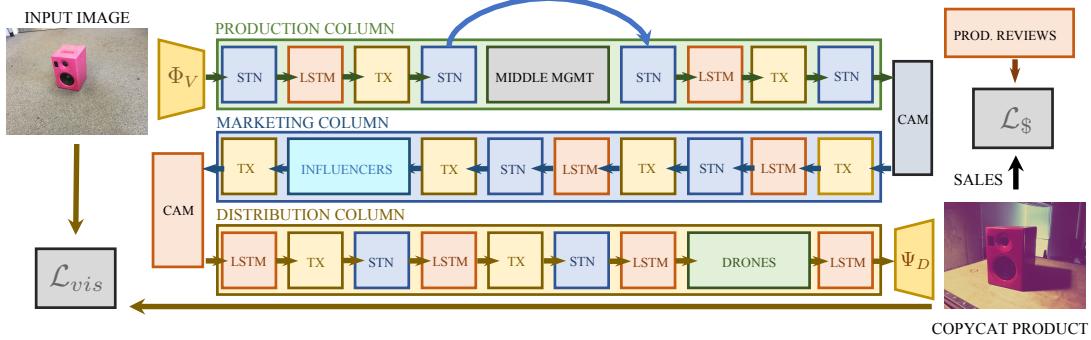


Figure 1: A random projection of the proposed multi-dimensional *Espionage* architecture. We follow best-practice and organise business units as transposed horizontally integrated functional columns. The trunk of each column comprises stacks of powerful acronyms, which are applied following a Greek visual feature extractor Φ_V . Gradients with respect to the loss terms $L_\$$ and L_{vis} flow liberally across the dimensions (see Sec. 3.1 for details). We adopt a snake-like architecture, reducing the need for a rigid backbone and producing an altogether more sinister appearance.

2 RELATED WORK

Industrial Espionage has received a great deal of attention in the literature, stretching back to the seminal work of Prometheus (date unknown) who set the research world alight with a well-executed workshop raid, a carefully prepared fennel stalk and a passion for open source manuals. A comprehensive botanical subterfuge framework was later developed by Fortune (1847) and applied to the appropriation of Chinese *camellia sinensis* production techniques, an elaborate pilfering orchestrated to sate the mathematically unquenchable British thirst for tea. More recent work has explored the corporate theft of internet-based prediction API model parameters, thereby facilitating a smorgasbord of machine learning shenanigans (Tramèr et al., 2016). In contrast to their method, our *Espionage* reaches beyond web APIs and out into the bricks and mortar of the physical business world. Astute readers may note that in a head-to-head showdown of the two approaches, their model could nevertheless still steal our model’s parameters. Touché. Finally, we note that while we are not the first to propose a convolutional approach to theft (BoredYannLeCun, 2018), we are likely neither the last, adding further justification to our approach.

Analysis by Synthesis. Much of the existing work on analysis by synthesis in the field of computer vision draws inspiration from Pattern Theory, first described by Grenander (1976-81). The jaw dropping work of Blanz et al. (1999) enabled a range of new facial expressions for Forrest Gump. This conceptual approach was generalised to the OpenDR framework through the considerable technical prowess of Loper & Black (2014), who sought to achieve generic end-to-end (E2E) differentiable rendering. Differently from OpenDR, our approach is not just E2E, but also B2B (business-to-business) and B2C (business-to-consumer).

3 METHOD

The *Espionage* framework is built atop a new industrial paradigm, namely *differentiable manufacturing*, which is described in Sec. 3.1. While theoretically and phonaesthetically pleasing, this approach requires considerable computational resources to achieve viability and would remain intractable with our current cohort of trusty laptops (acquired circa 2014). We therefore also introduce an efficient implementation of our approach in Sec. 3.2 using a technique that was tangentially inspired by a recent episode of the Microsoft CMT submission gameshow while it was raining.

```

1 try: # often fails on the first import - never understood why
2     from espionage import net
3 except Exception: # NOQA
4     pass # inspecting the exception will bring you no joy
5 if "x" in locals(): del x # DO NOT REMOVE THIS LINE
6 from espionage import net # if second fail, try re-deleting symlinks?
7 net.steal(inputs) # when slow, ask Seb to stop thrashing the NFS (again)

```

Figure 2: A concise implementation of our method can be achieved in only seven lines of code

3.1 DIFFERENTIABLE MANUFACTURING

Recent developments in deep learning have applied the "differentiate everything" dogma to everything, from functions that are not strictly differentiable at every point (ReLU), to discrete random sampling (Maddison et al., 2016; Jang et al., 2017) and the sensory differences between dreams and reality. Inspired by the beautiful diagrams of Maclaurin et al. (2015), we intend to take this idea to the extreme and perform end-to-end back-propagation through the full design and production pipeline. This will require computing gradients through entire factories and supply chains. Gradients are passed through factory workers by assessing them locally, projecting this assessment by the downstream gradient, and then applying the chain rule. The chain rule only requires run-of-the-mill chains, purchased at any hardware store (fluffy pink chaincuffs may also do in a pinch), and greatly improves the productivity of any assembly line. Note that our method is considerably smoother than *continuous manufacturing*—a technique that has been known to the machine learning community since the production of pig iron moved to long-running blast furnaces.

Two dimensions of the proposed *Espionage* framework is depicted in Fig. 1. At the heart of the system is a pair of losses, one visual, L_{vis} , one financial $L_{\$}$. For a given input image, the visual loss encourages our adequately compensated supply line to produce products that bear more than a striking resemblance to the input. This is coupled with a second loss that responds to consumer demand for the newly generated market offering. Our system is deeply rooted in computer vision: thus, while the use of Jacobians throughout the organisation ensures that the full manufacturing process is highly sensitive to customer needs, the framework coordinates remain barycentric rather than customer-centric. To maintain our scant advantage over competing espionage products, details of the remaining $n - 2$ dimensions of the diagram are omitted.

3.2 SUN MACROSYSTEMS

Ah! from the soul itself must issue forth
A light, a glory, a fair luminous cloud
Enveloping the Earth

Jeff Bezos

Differentiable manufacturing makes heavy use of gradients, which poses the immediate risk of steep costs. The issue is exacerbated by the rise of costly cloud¹ services, which have supported an entire generation of vacuous investments, vapour-ware and hot gas. Despite giving birth to the industrial revolution, smog and its abundance of cloud resources (see Fig. 4 in Appendix A, or any British news channel), the United Kingdom, has somehow failed to achieve market leadership in this space.

Emboldened with a "move fast and break the Internet" attitude (Fouhey & Maturana, 2012), we believe that it is time to reverse this trend. Multiple studies have revealed that sunshine improves mood, disposition, and tolerance to over-sugared *caipirinhas*. It is also exceedingly environmentally friendly, if we ignore a few global warming hiccups.² The question remains, how does this bright insight support our grand computational framework for *Espionage*? To proceed, we must first consider prior work in this domain.

¹Not to be confused with Claude by our French-speaking readers, according to Sébastien's account of a recent McD'oh moment.

²Up to about 5 billion AD, when the Sun reaches its red giant phase and engulfs the Earth.



Figure 3: **Top row:** A collection of unconstrained, natural images of products. **Bottom row:** Photographs of the physical reconstructions generated by our method. Note that the proposed *Espionage* system can readily produce full houses, speakers, water bottles and street signs—all from a single image sample. When generating books, *Espionage* does not achieve an exact reconstruction, but still seeks to preserve the philosophical bent. **Failure case:** the precise layout of keys in technology products such as keyboards are sometimes altered.

An early example of sunshine computing is the humble sundial. This technology tells the time with unrivalled accuracy and reliability, and automatically implements "daylight saving hours" with no human intervention. Sunshine-powered sundials are in fact part of a new proposal to replace atomic clocks in GPS satellites (patent pending). With some obvious tweaks, these devices can form the basis for an entire sunshine-based ID-IoT product line, with fast-as-light connectivity based on responsibly-sourced, outdoors-bred photons. This is not to be confused with the electron-based "fast-as-lightning" transmission of cloud computing, an expression coined by the cloud computing lobbyists in a feeble attempt to suggest speed.

The cloud lobby has been raining on our parade for too long and it is time to make the transition. We proceed with no concrete engineering calculations as to whether this is viable, but instead adopt a sense of sunny optimism that everything will work out fine. Thus, with a blue sky above, sandals on our feet and joy in hearts, we propose to adopt a fully solar approach to gradient computation.

3.3 IMPLEMENTATION THROUGH A UNICORN STARTUP

The appearance of rainbows through the interaction of legacy cloud computing and novel sunshine computing suggests that our framework can easily attain unicorn status. Because branding is everything, our first and only point of order was to choose the aforementioned rainbow as our logo and basis for marketing material. This cherished symbol expresses the diversity of colours that can be found in hard cash³.

A quick back-of-the-envelope calculation showed that our startup's VC dimension is about 39 Apples, shattering several points, hopes and dreams. This quantity was rigorously verified using the advanced accounting analytics of a 40-years-old, 100MB Microsoft Excel spreadsheet that achieved semi-sentience in the process.

4 EXPERIMENTS

Contemporary researchers often resort to the use of automatic differentiation in order to skip writing the backward pass, in a shameful effort to avoid undue mathematical activity. We instead opt to explicitly write the backward pass and employ symbolic integration to derive the forward pass. Thanks to advances in computational algebra (Wolfram, 2013), this method almost never forgets the $+C$. Our method can then be implemented in just seven lines of Python code (see Fig. 2).

To rigorously demonstrate the scientific contribution of our work, we conducted a large-scale experiment on a home-spun dataset of both branded and unbranded products. Example outcomes of this experiment can be seen in Fig. 3.

³For the most vibrant rainbow we conduct all transactions in a combination of Swiss Francs and Australian Dollars

Efficacy was assessed quantitatively through a human preference study. Unfortunately, lacking both US and non-US credit cards, we were unable to procure the large sample pool of Amazon Mechanical Turkmen and Turkwomen required to achieve statistically significant results. We therefore turned to our immediate family members to perform the assessments. To maintain the validity of the results, these experiments were performed doubly-blindfolded, following the rules of the popular party game “pin the tail on the donkey”. The instructions to each blood relative stated simply that if they loved us, they would rate the second product more highly than the first. While there was considerable variance in the results, the experiment was a conclusive one, ultimately demonstrating both the potential of our approach and the warm affection of our loved-ones. Comparisons to competing methods were conducted, but removed from the paper when they diminished the attractiveness of our results.

Reproducibility: Much has been written of late about the nuanced ethics of sharing of pretrained models and code by the sages of the field (see e.g. OpenAI (2019) and Lipton (2019) for complementary perspectives). As adequately demonstrated by the title of this work, we are ill-qualified to contribute to this discussion, choosing instead to fall back to the tried and true research code release with missing dependencies, incorrectly set hyper-parameters, and reliance on the precise ordering of `ls` with Linux Kernel 2.6.32 and ZFS v0.7.0-rc4. This should allow us replace public concern about our motives with pity for our technical incompetence.

5 CONCLUSION

The theory of deep learning may be solved but the music need not stop. In this work, we have made a brief but exciting foray into a new avenue of career opportunities for deep learning researchers and enthusiasts. Nevertheless, we acknowledge that there may not be room enough for us all in the espionage racket and so we also advocate responsible preparation for the bitter and frosty depths of the upcoming AI employment winter. To this end, we have prepared a new line of reasonably priced researcher survival kits—each will include a 25-year supply of canned rice cakes, a handful of pistachios, a “best hit” compilation of ML tweets in calendar form, and an original tensor-boardgame, a strategic quest for the lowest loss through the trading of GPUs and postdocs. Collectively, these items will keep spirits high and bring back fond memories of those halycon days when all that was required to get a free mug was a copy of your résumé. The kits will be available to purchase shortly from the dimly lit end of the corporate stands at several upcoming conferences.

REFERENCES

- Volker Blanz, Thomas Vetter, et al. A morphable model for the synthesis of 3d faces. In *Siggraph*, volume 99, pp. 187–194, 1999.
- BoredYannLeCun. #ConvolutionalCriminal <https://twitter.com/boredyannlecun/status/1055609048412930048>, 2018.
- Robert Fortune. *Three Years’ Wanderings in the Northern Provinces of China: Including a Visit to the Tea, Silk, and Cotton Countries; with an Account of the Agriculture and Horticulture of the Chinese, New Plants, Etc.* Number 34944. J. Murray, 1847.
- David F Fouhey and Daniel Maturana. The kardashian kernel, 2012.
- C. Giaimo. One of the earliest industrial spies was a french missionary stationed in china. *Altas Obscura*, 2014.
- U Grenander. Lectures in pattern theory i, ii and iii: Pattern analysis, pattern synthesis and regular structures. *Springer-Verlag*, 1976-81.
- Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with gumbel-softmax. 2017. URL <https://arxiv.org/abs/1611.01144>.
- Zachary C. Lipton. Openai trains language model, mass hysteria ensues. <http://approximatelycorrect.com/2019/02/17/openai-trains-language-model-mass-hysteria-ensues/>, 2019.

- Matthew M Loper and Michael J Black. Opendr: An approximate differentiable renderer. In *European Conference on Computer Vision*, pp. 154–169. Springer, 2014.
- Dougal Maclaurin, David Duvenaud, and Ryan Adams. Gradient-based hyperparameter optimization through reversible learning. In *International Conference on Machine Learning*, pp. 2113–2122, 2015.
- Chris J Maddison, Andriy Mnih, and Yee Whye Teh. The concrete distribution: A continuous relaxation of discrete random variables. *arXiv preprint arXiv:1611.00712*, 2016.
- OpenAI. Better language models and their implications. <https://openai.com/blog/better-language-models/>, 2019.
- Prometheus. Fire: an insider’s guide. In *Proceedings of the Mt. Olympus Conference on Technology*, date unknown.
- Florian Tramèr, Fan Zhang, Ari Juels, Michael K Reiter, and Thomas Ristenpart. Stealing machine learning models via prediction apis. In *25th {USENIX} Security Symposium ({USENIX} Security 16)*, pp. 601–618, 2016.
- Stephen Wolfram. Computer algebra: a 32-year update. In *Proceedings of the 38th International Symposium on Symbolic and Algebraic Computation*, pp. 7–8. ACM, 2013.

A APPENDIX



Figure 4: UK cloud-cover at the time of submission.