The Affective Growth of Computer Vision

Norman Makoto Su David J. Crandall
Luddy School of Informatics, Computing, and Engineering
Indiana University Bloomington
{normsu,djcran}@indiana.edu

Abstract

The success of deep learning has led to intense growth and interest in computer vision, along with concerns about its potential impact on society. Yet we know little about how these changes have affected the people that research and practice computer vision: we as a community spend so much effort trying to replicate the abilities of humans, but so little time considering the impact of this work on ourselves. In this paper, we report on a study in which we asked computer vision researchers and practitioners to write stories about emotionally-salient events that happened to them. Our analysis of over 50 responses found tremendous affective (emotional) strain in the computer vision community. While many describe excitement and success, we found strikingly frequent feelings of isolation, cynicism, apathy, and exasperation over the state of the field. This is especially true among people who do not share the unbridled enthusiasm for normative standards for computer vision research and who do not see themselves as part of the "in-crowd." Our findings suggest that these feelings are closely tied to the kinds of research and professional practices now expected in computer vision. We argue that as a community with significant stature, we need to work towards an inclusive culture that makes transparent and addresses the real emotional toil of its members.

1. Introduction

By almost any metric, computer vision is in a golden age. Deep learning has revolutionized nearly every problem in computer vision, achieving results that were unimaginable just a few years ago. As technical barriers have fallen, the floodgates of interest in computer vision have opened wide [40]. CVPR 2019 had over 9,000 attendees – nearly ten times more than CVPR 2009 – and paper submissions were growing exponentially, leading the 2019 Program Chairs to cheekily extrapolate that CVPR 2028 will surpass 10 billion submissions [5].

Meanwhile, the rest of the world is paying attention to computer vision. According to Google Scholar, CVPR is the fifth most impactful publication venue of all journals and conferences *in all of science* [2], just below *The Lancet* and significantly above *PNAS*. Governments around the world have announced major investments in AI research over the next decade [14, 37], and academic grants are increasingly focused on proposals integrating machine learning, sometimes with industry backing [3]. Major tech companies are investing billions of dollars into computer vision and machine learning [4] and paying eye-popping salaries to attract top talent [32].

Despite all this apparent success, it is clear that computer vision is facing a number of important challenges. The thousands of papers published every year in computer vision – and the popularity of arXiv as an instant publication venue – have made it impossible for any individual to follow all developments in the field. Peer review processes are straining under the huge influx of paper submissions and the limited number of qualified reviewers, leading to acceptance decisions that may often be arbitrary [26]. Faculty spending time in industry may harm student success [33], but those staying in academia find it increasingly difficult to compute against the computational, data, and human resources of industry labs. Researchers in industry enjoy much of the freedom of academia, but without the formal protections of tenure [34]. Meanwhile, there is a striking lack of diversity across gender and race in AI [49].

There is also growing alarm about the ethical consequences of computer vision, from self-driving cars that have killed people due to perception failures [22], to surveillance technology that could become a tool of repression [18], to the exploitation of low-income workers for labeling training data [50], to face recognition algorithms that exhibit racial bias [12]. Major datasets that have been the bedrock of computer vision research for a decade have been found to inadvertently include biased, racist, and misogynistic images and labels, leading the Tiny Images dataset to be formally retracted [47] and hundreds of thousands of images to be removed from ImageNet [44].

It is against this backdrop that we, the researchers and practitioners of the computer vision community, live and work. Given the field's implicit goal to build models that can match or beat human perceptual capabilities – to, arguably, "replace" humans – it is perhaps unsurprising that there has been little work studying the effect that this environment has had on those working in the field. The work of computer vision – just like any tech work [24, 31] – is laden with emotion: the joys of solving a problem, the sting of a paper rejection, the anxiety of finding a job or attaining tenure, the envy for others who are better known, the concern about ethical implications. These emotions orient how one behaves and drive the trajectory of the community [20].

In this paper, we set out to investigate the "affects" (emotions) of working in the modern vision community. We invited several hundred computer vision and machine learning researchers to participate in a study in which we asked them to write a short story about a recent, specific incident involving their work that was emotionally salient to them. We collected and analyzed 56 stories from a wide range of informants, ranging from first-year Ph.D. students to prominent senior researchers in academia and industry. While many report being excited about the progress in the field, we were struck by the number of stories indicating nostalgia about the past and worry about where the field is going. Many informants seem to feel isolated and are struggling to find their place within the field. Some were so nervous about the consequences of sharing their stories that they were only willing to communicate with us anonymously through a third party.

We are well aware that this is not a typical CVPR paper: we beat no benchmarks, we introduce no datasets, we present no novel loss functions. But we argue that this type of work is nevertheless of central importance to CVPR, not to be sidelined off its main track. We identify how the growth of computer vision has affected a diverse range of individuals; these affects cannot be measured in terms of aggregate quantitative metrics. This paper seeks to amplify these emotions, with the goal of taking a first step towards a longer-term conversation about where our community is going and how to ensure that it is vibrant for all.

2. Related Work

Our paper is reminiscent of work published at CVPR and other AI conferences that has not introduced technical innovation but instead studied the research community itself. Torralba and Efros [46] pointed out issues of bias in computer vision datasets back in 2011, while more recent work revealed systematic ethical issues with widely-used computer vision datasets [9]. Buolamwini and Gebru [12] show that popular face recognition algorithms systematically discriminate based on gender and race. Wagstaff [48] argued that the machine learning community had lost its "connection to problems of import to the larger world of science and society" and proposed solutions. Tomkins et al. [45] analyzed the peer review process of a major computer science conference and uncovered distressing patterns of bias.

This paper also turns its lens onto the research community to study the effect of recent developments in the field. Studies of the data science profession [23, 25, 36, 39, 51] have investigated the contingent nature of its practices, but not the affective charge of such practices nor of changes in the discipline as a whole. Perhaps the closest work to ours are podcasts and newsletters that have collected informal stories from prominent vision researchers [1, 38]. In contrast, we systematically study (see [41, p.19-22] for a distinction between journalism and social research) the affects of the computer vision community by soliciting anonymous stories from a wide range of researchers whose stories have not yet been told.

Critical scholars [8, 15] have investigated how structural biases in society are perpetuated by the datasets, algorithms, and stakeholders of AI. For example, face recognition is of large concern to marginalized groups [21] and performs worse in identifying such individuals [12, 42]. Machine learning can benefit members of dominant cultures, often at the cost of harming others, even if unintended [8, 15]. We build upon this important work here, centering on emotions of the computer vision community and identifying how growth of the discipline has affected those who research and practice computer vision and machine learning.

3. Methodology

We sought to collect diverse stories describing the emotional impact of working in computer vision and machine learning. Stories offer a way to access events that researchers cannot witness and to learn about their informants' experiences [29]. Informants were asked to write their story in online documents (Google Docs) with the following prompt (summarized for brevity's sake): *Write a nonfiction story about computer vision and/or machine learning (at least 2 paragraphs) with yourself as the main character. The story should involve recent changes in the profession/discipline and depict an event of emotional impact on you.* Such solicitations empower people to write emotionally-charged but concrete experiences in their own voice [35]. We chose this method – analogous to an asynchronous, text-based semi-structured interview – instead of live interviews to facilitate flexible scheduling during the COVID-19 pandemic, to help informants feel comfortable sharing emotional stories, and to elicit a diversity of experiences [30]. After each informant drafted their story, we asked questions via Google Doc comments, and they then replied or updated the draft. This repeated until we had no further comments. This process emulates best practices for qualitative interviewing [16].

Informants also filled out a short demographic survey and completed a group identification scale [10] to measure the degree to which they identify with computer vision (or their other stated home field). This scale gives a rough sense for how closely the informant perceived themselves as part of the "in-group" of their field.

We conducted the study from May through November 2020. Participants' identities were kept confidential, and the study was approved by our university's ethics board. We advertised using a variety of means: (1) placing an ad in the PAMI-TC and IAPR email newsletters, (2) sending direct emails to about 100 computer vision and machine learning PIs and asking them to distribute to their labs, (3) sending direct emails to about 100 students, postdocs, and industry researchers, (4) contacting organizers of affinity groups Queer in AI, LatinX in AI, Black in AI, and Women in Computer Vision and asking them to share with their email lists, (5) posting on social media, including to the "Computer Vision" and "Computer Vision and Image Processing" groups on Facebook, (6) making announcements in virtual rooms of ECCV 2020, and (7) via snowball sampling [16].

Whereas quantitative approaches test *a priori* hypotheses about the distribution of a population among known categories and develop predictive models, our qualitative analysis sought to *discover* relevant categories related to phenomena [28] – the emotional responses to computer vision's explosive growth. We followed a constructivist grounded theory approach [13] to analysis, employing initial incident-level codes that included, for example, "awe with deep learning," "unease with the blackbox," and "feeling left behind the hype." Focused and axial coding further developed themes around emotions tied to the new prestige of computer vision, shifting norms of science, increased prominence of industry, possible harms of AI, and rise in celebrity culture. Memos on these themes form the genesis of our findings.

4. Findings

In total, 103 people responded to our call for participants, 64 completed the demographic survey, and 56 completed the entire study. As shown in Table 1, our informants represented a diverse sample of the computer vision community, roughly evenly split between academia and industry, and at a range of levels of seniority from Ph.D. students through senior scientists and full professors. In terms of location, about 80% of our informants were currently in North America, about 15% were in Europe, and only about 5% were in Asia. Compared to attendance at CVPR 2019 (56.2% North America, 13.7% Europe, 28.5% Asia, 1.0% Oceania, 0.5% South America, 0.07% Africa) [5], our sample significantly under-represents Asia, and thus our findings best capture sentiments of researchers in North America and Europe. The median age of informants in our sample was 33 years ($\mu = 35.3, \sigma = 9.4$), while the median number of years working in the field was 9.5 ($\mu = 10.6, \sigma = 7.6$). The majority (66.1%) of informants identified their primary field as computer vision, while 21.4% identified artificial intelligence, 7.1% identified machine learning, 3.6% identified natural language processing, and 3.6% identified another area. According to the group identification scale, informants on average reported being on the high end of belonging to their primary field ($\mu = 41.0, \sigma = 6.6$ out of a maximum possible of 50). We do not present individual demographic data because of the sensitivity of the topic and the possibility that such data could be triangulated to reveal identities [11]. To protect identities, all informants are referred to by an ID number. Block and italicized quotes are verbatim from stories.

Based on our qualitative analysis, we reveal experiences running the gamut of human emotions. On the positive end, deep learning (DL) has transformed computer vision into a field with immense real-world impact and prestige. Informants identify a sense of awe and delight with how DL has almost magically revolutionized nearly every problem in the field. About 21% (12/56) of our stories expressed solely positive feelings. The majority of stories, however, paint a general mood of malaise. Informants describe emotions of isolation, anger, apathy, and cynicism regarding the growth of the discipline (and its side effects). These feelings are tied to frustration over the loss of "science" as well as to the increasingly competitive nature of the community.

4.1. The Magic of Deep Learning (DL)

Stories depict the resurgence of deep learning around 2012 as suddenly paving a way to making computer vision work in real-world applications. P21 described feeling that it was a *"tough and hopeless time"* in computer vision *"before 2012, [when] the annual performance improvements over ImageNet are quite marginal."* Some informants describe initially feeling skeptical about DL's potential. For example, P5 describes feeling *"shocked"* that a colleague told him to abandon the use of a conditional random field for his problem: *"she told me you should solve the problem purely based on deep learning... I did not think the occlusion problem can be solved without explicitly reasoning of shape priors and depth ordering."*

Of course, convolutional neural networks for DL turned out to be a *"hammer"* (P5) that produced undeniably good results. P8 describes amazement in how he replaced a probabilistic graphical model with a new deep neutral network he devised, which was *"6 times faster... and... applicable to large-scale data."* He felt pride in making a model that is now used day-to-day in his application area.

The speed and ease with which these results could be achieved was also astounding. P1 revisited an old paper of his advisor's, reimplemented it via deep learning, and within half an hour it *"outperformed the prior model by*

Gender		Ethnicity		Highest education		Current location		Current employer		Current job title	
Male	73.2%	Asian	44.7%	Ph.D.	73.2%	North America	80.4%	Research University	58.9%	Assistant Professor	26.8%
Female	23.2%	White	37.5%	Masters	14.3%	Europe	14.3%	Industry	28.5%	Ph.D. student	16.1%
Non-binary	3.6%	Hispanic	3.6%	Bachelors	12.5%	Asia	1.8%	Other academia	7.1%	Scientist/Engineer	26.7%
		Middle Eastern	3.6%			Australia/NZ	1.8%	Start-up	3.6%	Associate Professor	8.9%
		Black	1.8%			Decline to say	1.8%	Government	5.4%	Professor	8.9%
		Decline to say	10.7%					Decline to say	1.8%	Postdoc	5.4%
										Sr Scientist/Engineer	8.9%
										Instructor	1.8%

Table 1. Demographics of informants ($N = 56$). Some columns do not sum to 100% because informants identified with multiple options.

10–20%. I remember bouncing back into my advisor's office with a silly grin." The way in which DL seemed to magically work on so many problems also made informants reflective, thinking of their previous efforts in *"feature engineering, clustering, and classifier design"* (P1) that now seemed like dead ends. P51, after beating his previous model after half a day's work, felt *"ashamed. . . that my solution was so obsolete. Sometimes I wonder whether it was my fault (I was not up-to-date enough) or not (state-of-the-art changed quickly)."* Shame is bound up with fears of being left behind, which we will discuss later.

4.2. Frustration over Shifting Scholarship

Many stories expressed a general sense that DL has shifted, in a number of ways, the meaning of "science" in the discipline. First, informants lamented how DL has made computer vision into more of an engineering exercise. Second, stories express fear that one may become responsible for creating problematic systems, a direct result of this engineering drive. Last, informants complained of a collective laser-like focus on DL that has created a selective amnesia that rewards some ways of approaching problems while discouraging others.

4.2.1 From Scientists to Neural Network Technicians

For some informants, what attracted them to their disciplines was the dream of shedding light onto fundamental questions about, for example, how humans see. P7 notes *"deep learning-based systems are trained for very specific objectives, and are far from resembling anything that could be considered a general model."* What is lost for P7 is the original ambitious goal of computer vision to develop a general model that could solve many different everyday tasks, one that reaches an understanding of the *"mechanisms of perception"* and does not rely on tech companies' agendas of *"big data from the Internet."* P14 came to computer vision to understand the *"fundamental question of how biological vision works."* For others, computer vision held the promise of understanding how we think and learn – *"I want to understand the learning process, and this SOTA [state-of-the-art] chasing isn't that"* (P23).

Of course, computer vision has a long history of investigating approaches that work well in practice but are not biologically motivated; in fact, convolutional neural networks are probably more biologically plausible [7] than the approaches based on SVMs with SIFT or HOG feature vectors [19,43] that were popular before DL's resurgence. Nevertheless, many informants felt that DL's dominance has shifted the field from a focus on what they consider to be the fundamentals. P15, who was trained in cognitive science, observes a shift from people wanting to build models that *"*explain* the internal cognition of people"* to those that merely *"[describe] the external behavior of people,"* and shares a particular example of asking a class how they would build AI for a board game called Dixit:

> I thought the students would start talking about all the. . . intelligence that go into playing Dixit, from really really complex visual recognition and interpretation of the beautiful and surreal Dixit artwork cards, and the emotions and moods that each card conveys, and referencing cultural knowledge and commonsense knowledge, and having to. . . model something about the other players, including things you know about them personally, like even inside jokes you might have had with them from long ago, and creative linguistic expression. . . and I thought. . . students would realize how complex and mysterious is the human mind. . . and how in AI we have barely scraped the surface of all of these wonderful mysteries, and they would all go away from that class totally mesmerized by. . . this insightful activity. . . that I had masterfully orchestrated for them. . .
>
> Instead. . . they (excitedly, and confidently) said. . . "You could collect a large amount of data from people playing Dixit, and then train a neural network to give responses to cards!". . . I think I stood there for a few seconds somewhat taken aback and not quite sure how to respond to this confident answer chorus. My mouth was probably hanging open. . . The students were convinced. Dixit was, after all, a very easy problem to solve. Nothing to see here. Move along, move along. Another day in AI. Another win for neural networks.

P15's point is that students are now stuck in this deep-learning mode of thought, unable to consider other approaches. This narrow perspective – and a perceived focus on beating benchmarks as opposed to advancing science – was a recurring theme of informants' stories. For example, P49 told of a paper submission that showed that a state-of-the-art model had overfit to the test dataset, and that their own simpler approach generalized better with much less training data. But it was rejected because it did not, ironically, beat the state-of-the-art on the original, overfit dataset. The reviews made no mention of the key argument: *"criticism of a well-known dataset."*

This shift is belittled by some as transforming what was once a scientific discipline into one that emphasizes *"solution method[s]"* (P15) and *"engineering the black box"* (P46). P25 remembers being outraged at a talk in which the speaker concluded by saying, *"Remember, we are neural network technicians, not scientists."* Overall, stories spoke of much work that consists primarily of tweaking parameters and architectures to make incremental progress on shallow metrics. Informants bemoaned *"how much [computational] power... are 'wasted' just to get that less than 1%... improvement on the accuracy"* (P13), felt silly and unenthusiastic over gaining *"1% accuracy on imagenet accuracy"* (P23), and were demotivated because they are forced to do "research" for *"established (un-novel) deliverables, e.g., running known methods on their data"* (P28).

This new way of doing science is also puzzling to those outside of machine learning. P39 related a story about a collaboration with domain scientists:

> My student had developed an initial 2D CNN model which was doing reasonably well on the difficult problem... My student and I were describing the large number of decision choices in developing such a model such as the number of convolutional layers, the size and number of the kernels in each layer, etc. Our collaborators, the domain scientists, asked if we had done an in-depth study on the optimal choice of each of these parameters. They were accustomed to more traditional, simpler models in which it is feasible to do possibly exhaustive search on the optimal parameter settings. However, we had to explain to our collaborators that it really isn't feasible to do such a search – there are just too many parameters and other design choices. I could sense that this was somewhat disillusioning for them, not knowing whether the model was the optimal one.

While the collaborators came to accept DL, they needed a lesson on the immense complexity of DL models – and the difficulty of understanding what they are actually modeling.

4.2.2 Paranoia and Fatigue over Harmful Blackboxes

As computer vision's reach in our everyday lives expands, stories spoke of a growing realization among researchers that they could no longer be a *"simple happy nerd pushing boundaries on the next cool computer vision technology,"* or *"pretend to be an ostrich researcher hiding my head in the sand and blame others for the misuse of technology"* (P18).

P42 felt a growing paranoia about his role in building an AI system for a company *"because the system is so much of a black box, trying to build in explainability and transparency into the system feels inherently futile sometimes. The most we can do is really focus on what inputs are going into the system, weights, and training data. There's focus on system design so we aren't inadvertently presenting the output in a way that either obfuscates the confidence/assumptions of the system and also doesn't portray the system as 'all-knowing' or too cautious/gung-ho. It makes me feel better but I hope the people using this system understand it's not some magic box."* P42 is conveying a feeling, exacerbated by DL's blackbox nature, of uncertainty regarding the social ramifications of his system [23].

Some stories depict informants disgusted at the *"gaslighting and undervaluing"* (P23) of critical AI research. P41, who is transgender, describes their experience reading a paper on facial gender recognition that motivates its system for identifying transgender people by the specious claim *"that some bad actors could be taking HRT [hormone replacement therapy] as a disguise technique to spoof face recognition algorithms."* They felt *"pure rage as first reaction, and then just a deep sadness, which still persists. And it piles up on a stack of research being done... that I consider wrong and/or unethical, and made me lose any excitement for the field."* They note the authors apologized, but that the problem is much deeper than one single paper. *"There are not many ML projects that I read about these days that I think should exist and I am aware of the weight of this statement... I am tired these days."*

The mindset of applying DL on big data to solve problems, without necessarily stopping to think of the consequences, may in part be driven by intense pressure to publish. P46, a senior faculty member, commented, *"Having something accepted appears to be more important than having something good accepted."*

4.2.3 Selective Amnesia

Closely intertwined with the shift described above is an effective erasure of past work. P7 calls this *"selective amnesia."* P19 tells of feeling helpless when students are unable to comprehend a classic paper *"on spatio-temporal interest points... The student spent more than a week and returned completely puzzled by this paper written in the dinosaurs era"* before deep learning. P48 is *"often sad that I am supposed to fetishize particular techniques just because they are new. I don't like shiny/impressive stuff, I like thorough stuff! Nuanced stuff!"*

This disregard for older work also appears in conferences. When speaking with a poster presenter whose work is related to their own done five years ago, P19 is deflated with the response: *"Heh, I don't read any papers before*

2015." Fundamentally, P19 – a senior faculty member – worries that *"we are a generation that will forever be blamed for breaking the flow of science."* Here we see not only a dismissal of prior work, but also a surprising disrespect for experienced researchers themselves.

4.3. Newfound Opportunities

Aside from research, informants identify the newfound applicability of computer vision and its related disciplines as affecting their professional and personal lives in positive ways. P11 describes that years ago before the current AI boom, the public doubted the *"practical effects"* of computer vision. His future father-in-law interrogated him about the future of the field: *"While well prepared, I still suffered... After I left their home, my girlfriend's father... expressed his concern against our relationship. He thought 'it is hard for me in this field to find a good job in the US, even in China.'"* He felt sad at the time, but he reported a happy ending: *"our surroundings currently are full of news on computer vision and artificial intelligence applications."* Computer vision is now well-known to the public, and its job prospects are bright.

The rapid rise in importance of industrial research labs surprised many informants, especially faculty. P38 remembers a keynote speaker *"asked people to raise their hand if they were in industry, and ... about half of the people raised their hand, which was really surprising to me at the time."* P34 recalls *"very few industry positions"* available in 2012 with only 20 companies at CVPR; when he graduated in 2017, he felt excited that *"hundreds [of] companies... showed up at our conferences."*

As a PI in computer vision for almost a decade, P4 feels envy: *"[students] don't even realize how much better they have it."* She recalled receiving an email from a well known researcher, "Yusuf" – who has made fundamental scientific contributions – regarding her PhD student, "Sanjeev:" *"[Yusuf] had come across Sanjeev's paper... and wanted to know if Sanjeev might be interested in an internship over the upcoming summer. Sanjeev has spoken to Yusuf on the phone. Sanjeev tells me that while the conversation was pleasant, he is not all that excited about the internship."* This shocked P4; when she was a student, internships were rare, senior researchers did not cold-email students, and students did not dismiss such opportunities so casually. Here is a dynamic switch in power.

Lastly, stories comment on the now carnival-like and opulent character of conferences – *"lavish corporate diners... with the usual flowing wine, well appointed buffet, and irresistible pastries"* (P7). P33's story vividly captures the CVPR experience by describing its host of characters from poster presenters who *"carry tubes... like... martial artists carry their swords to attend an annual grand tournament,"* a grizzled professor whose contributions belong to a *"prehistoric geometry-heavy era,"* two old colleagues reminiscing *"when CVPR was with two hundred attendees,"* young scholars who keep socializing at the *"after-hour parties... sponsored by a ride-sharing/auto-driving NASDAQ listed company... with bands rocking every participant,"* and a professor *"showing off his lab's technology adopted on a live-streaming App that makes billions."*

4.4. Industry Reinforcing the New DL Science

Stories also commented on the downsides of the increasing prominence of industry. P28 describes how many companies are eager to jump on the AI hype train: "*Industry is especially excited to be involved, as everyone wants to advertise that their product uses A.I. and is therefore faster, and smarter, than the products of their competitors. So, there has been an explosion in the number of businesses with people who don't understand machine learning, but who need researchers to revolutionize their product.*" P28's PI has capitalized on this, creating a lab structure where everyone *"is assigned to work at a company for funding. And these companies make us work *hard*. Weekly deadlines and everyday meetings are the norm. This would be fine if what we were doing was research, but it is not. These companies do not have the time to let us... explore novel methods."* P28 describes students as feeling angry and demoralized because they essentially do industry work on graduate student salaries.

The job market is so competitive that we hear stories of students who successfully sell themselves as DL experts, despite being ill-prepared. P20 complains that half of their job candidates *"cannot calculate the output dimension correctly given input dimension and kernel dimension; some tries to implement convolution with standard definition from math... which makes the implementation complicated and less efficient (with more cache miss because one of the arrays needs to be accessed backwards); some struggles when trying to handle the case of kernel stride* $>$ *1."* P47 is shocked to hear a mediocre M.S. student confidently declare they would find a research position in industry, only to discover a few weeks later they were hired as a senior research scientist. A year later, P47 runs into the student at CVPR: *"I learned he was still in the same startup... However, he had some issues... trying to calibrate some cameras, not sure how to compile some software library, unsure what type of capture setup was adequate for his problem... The conversation turned into just another of our old office meetings where it was clear he was completely lost and in need of... hand-holding to get out of the rut he was in."*

4.5. Not a Cool Kid

Some of the most emotionally-impactful stories come from those who feel left behind; in fact, we were surprised at how widespread this feeling was and how it came even from faculty and researchers at top labs. It is manifested in multiple ways. First, conferences seem to be less and less conducive for junior and senior researchers to network. P26, a faculty member at a top university, used to believe that good work would naturally be recognized but was stunned to discover that a senior colleague who had written letters for him in the past was unaware of his recent work. P26 noted that *"the sheer volume of papers... means that few people even look at all the papers in their area... Workshops have talks, but... [prefer] instead to call upon the same set of senior researchers again and again."* Conferences encourage a celebrity culture; P3, a postdoc, talks of seeing senior faculty and feeling *"like a teenager is after a movie star"* because *"the power of the idol is incredible" – "they were glowing in my eyes."*

Even senior colleagues describe feelings of insecurity as others are caught up in this idolization of researchers. P29, another faculty member, describes feeling like a *"two-bit player in a giant circus"* at CVPR when a stranger approached his junior colleague for a selfie to post on social media: *"I realized I was more of a dinosaur than I thought. This colleague, my 'junior,' had had their research become a cause not just for science, but also for celebrity. I did not envy the colleague's fame, as much as I envied the fact that they belonged at CVPR more than I ever would."*

Stories spoke nostalgically of when CVPR was smaller and anyone could be heard. During his first CVPR a decade ago, P53, now an assistant professor, *"stayed at a 1-bedroom apartment in Chinatown (30 minutes walk to the venue) with four other students... I enjoyed every moment of the conference with the mindset of amateurs, attending all paper sessions, poster sessions, and social events hosted by the conference. I actively made many friends... (believe me, I am not that social person.), and with my very limited knowledge and technical skills."*

But the rapid changes in the field, the exponential growth of the conference, and the competitive environment have made CVPR seem unwelcoming for some outside its norm. P25, a professor at a top undergraduate teaching college, describes experiencing *"heartbreak"* at a recent CVPR. Though he attended many times before as a graduate student, he felt out of place trying to return to the conference after a few years gap: *"I tried to talk to presenters to get more insights about their work. But it was clear to... them... that I didn't have any insights to offer in return... I had already fallen behind. So they moved on and talked to others... I felt like the vision community had no place... for non-publishing undergraduate educators, even those who were preparing their future graduate students. So I withdrew. And I haven't returned to CVPR since then."*

Even those who successfully built their careers on DL found it hard to keep riding its wave. P2 felt he was at the right place at the right time, as one of the first to realize that *"deep learning would be the next big thing... I... told everyone in my lab about convolutional neural networks."* He published early work and *"people liked me for it."* He was successfully able to identify himself as part of a cadre of researchers doing DL. However, as DL became ubiquitous in computer vision, he was consumed by failure: *"I suppose it [DL] was an identity that had been stolen from me."* P2 fought to overcome these feelings through therapy. *"I am a person. I am also a researcher. I am not a Deep Learning researcher... That's enough for now."*

4.6. Marginalizing Ethics

Those concerned with the ethics of AI also feel excluded from mainstream research venues. P16, a Ph.D. student, excitedly attended NeurIPS for the first time and went to multiple affinity group workshops such as Women in Machine Learning and Queer in AI: *"The potential of artificial intelligence technologies to cause harm had been receiving increasing attention both from scientists and from the media... I felt that people were getting to the heart of the issues, talking about the role of power in artificial intelligence technologies. I left the conference that day feeling inspired and more excited about my field than I had felt in some time."* Yet after exuberantly telling her friend about the day's talks, she was surprised when he asked, *"Why are you wasting your time with these workshops instead of going to the core conference talks?"* These core topics were *"machine learning topics: optimization methods, network architectures, etc.,"* not *"diversity, equity, and justice."* The question hurt because she saw the affinity group conversations as vital to the field *"rather than as a side-show."* Yet, her friend disappointingly seemed to reflect the *"attitude of the field at large"* that these concerns are secondary.

5. Discussion and Conclusion

Our goal is quite unusual compared to most papers at CVPR: from the onset, we sought not to create new benchmarks in computer vision, but to benchmark the tenor of the computer vision research community. It is easy to take for granted how emotionally tied we are to the disciplines we belong to; certainly many of us have normalized the grit and suffering needed to succeed in research. The qualitative stories of researchers and practitioners we have collected and analyzed paint a complex, heterogeneous picture of the affective growth of the field. We do not claim one way of emoting is more legitimate or representative than another, but – just as Lipton and Steinhardt [27] surmise that the unprecedented growth of machine learning has led to "troubling trends" in scholarship – we see computer vision's meteoric rise as having an unprecedented emotional impact on its adherents. This paper identifies this *affective growth*, paving the way for future debate and actions. Are the sorts of feelings that people in computer vision are experiencing within the range we desire? If not, how could the CVPR community better support its individuals?

We summarize how deep learning (DL) developments have moved in tandem with increasing feelings of marginalization at different levels. Stories depict pride but also cynicism over the growing generational divide between those who were active in the pre-DL era and those who have only been exposed to DL-based techniques. There is undeniable euphoria with the breathtaking successes of DL. There are those that take great delight in engineering solutions, beating benchmarks, and finding real-world applications that the public will recognize. Yet many feel that computer vision has veered from the original goal that attracted them to it, away from fundamentals and towards engineering "black boxes." Others fear they are becoming "dinosaurs" whose expertise is no longer valued or relevant. There are those who remember when conferences were smaller, more egalitarian spaces where junior and senior faculty could readily network and exchange ideas. Students seem unable to break out of a solution-methods perspective. Some senior researchers, despite their status, feel marginalized.

Stories convey both joy and shock at the effects of industry's full-throated participation in the field. Whereas once computer vision was a sleepy corner of computer science, now students and faculty have countless opportunities as they are aggressively feted by industry. Academic conferences, now generously funded by industry, have started resembling trade shows or spectacles that some attendees love. Yet we hear of stories of problematic relationships between industry and research labs. Academic PIs, struggling to compete with the resources of industry, increasingly turn to industry joint appointments or collaborations. In the worst case, these collaborations can cause students to become demoralized as they work on projects that are driven by companies while the students are on graduate student salaries. In the rush to hire DL expertise, stories depict surprise at how under-prepared students can easily find jobs.

Stories convey paranoia, fatigue, rage, and isolation from those who do not fit the norm of computer vision researchers. Those adopting a more critical viewpoint of the technologies they are building have like-minded allies, but they feel such groups are treated as a sideshow to the real, "prestigious" work. Problems like face recognition that were viewed as innocuous technical challenges for decades suddenly seem problematic in the deep learning era, when the techniques work well enough to be applied to large-scale surveillance, and when their successes and failures have real consequences on real people's lives. Such problematic papers produce devastating emotional harm to their readers (such as the gender recognition paper mentioned above). Teaching-oriented faculty, vital to cultivating the next generation of graduate students, are "heartbroken" when they feel lost and undervalued in conferences.

Our goal is not to dictate whether or how the current landscape of emotions in computer vision should change, nor to offer prescriptions for how to do so. However, given the above landscape of emotions in computer vision, we do argue that special attention needs to be paid to individuals in the community who are experiencing significant – often negative – emotions that are not widely talked about. While the community has well-refined ways to guide its technical trajectory, through processes such as peer review, we are less well-equipped to monitor and make visible the emotions of the individuals in our community, and this may impede the trajectory toward healthy growth of the field. In the context of conferences, recognizing various feelings of marginalization may require rethinking the processes by which decisions can create more equitable relationships and opportunities. Data-driven instruments such as surveys to identify and act on concerns of the community, while ostensibly democratic, can end up perpetuating the status quo, the dominant view, rather than allowing the community to enable those who feel marginalized [17].

Machine learning and computer vision research communities have made recent changes to begin to address some of these concerns. CVPR 2019 introduced a Diversity Chair, for example, while authors at AAAI 2020 and NeurIPS 2020 were required to identify ethical consequences of their papers. CVPR 2020 issued a statement in support of the Black Lives Matter movement that called for a number of steps [6] including workshops and tutorials that examine "problems in equity, diversity and inclusion from a technical perspective" and "possible social damage flowing from computer vision technologies," and networking events and travel support to nurture "communities that are currently not well represented at CVPR." We envision our study – and others that build on it – will inform leadership groups like the PAMI Technical Committee and the Computer Vision Foundation about how the the changing tides of computer vision are affecting our community, and to consider measures to help chart an inclusive course for the future.

We close with the hope that this paper achieves its goals of uncovering and amplifying what some readers may already be feeling – perhaps in secret – about how the field they love has so transformed.

Acknowledgments. We thank Hamid Ekbia for discussions that inspired this paper, and Alexandra Branzan Albu, Arjan Kuijper, Maithilee Kunda, Linda O'Gorman, Devi Parikh, and Walter Scheirer for assistance with recruitment.

Disclosures. This work was funded entirely by Indiana University. The authors wish to disclose that they have received unrelated funding from DTRA, ETRI, Facebook, Google, IN3, NASA, NSF, ONR, and the U.S. Navy within the last five years. The views expressed in this paper are those of the authors and do not necessarily reflect the views of these or any other agencies.

References

[1] Computer Vision News - The Magazine of the Algorithm Community. https://www.rsipvision.com/computer-vision-news/. (accessed 2021-03-18). 2

[2] English - Google Scholar Metrics. https://scholar.google.com/citations?view_op=top_venues&hl=en. (accessed 2021-03-18). 1

[3] NSF Program on Fairness in Artificial Intelligence in Collaboration with Amazon — NSF - National Science Foundation. https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505651. (accessed 2020-10-01). 1

[4] OpenAI forms exclusive computing partnership with Microsoft to build new Azure AI supercomputing technologies. https://news.microsoft.com/2019/07/22/openai-forms-exclusive-computing-partnership-with-microsoft-to-build-new-azure-ai-supercomputing-technologies/, July 2019. (accessed 2021-03-28). 1

[5] Welcome to CVPR 2019. https://cvpr2019.thecvf.com/files/CVPR 2019 - Welcome Slides Final.pdf, 2019. (accessed 2021-03-18). 1, 3

[6] Inclusion Statement — CVPR2020. http://cvpr2020.thecvf.com/node/752, 2020. (accessed 2021-03-18). 8

[7] Sergey Bartunov, Adam Santoro, Blake A. Richards, Luke Marris, Geoffrey E. Hinton, and Timothy P. Lillicrap. Assessing the Scalability of Biologically-Motivated Deep Learning Algorithms and Architectures. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, pages 9390–9400, Red Hook, NY, USA, Dec. 2018. Curran Associates Inc. 4

[8] Ruha Benjamin. *Race after Technology: Abolitionist Tools for the New Jim Code*. Polity, Medford, MA, 2019. 2

[9] Abeba Birhane and Vinay Uday Prabhu. Large Image Datasets: A Pyrrhic Win for Computer Vision? In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1537–1547, 2021. https://openaccess.thecvf.com/content/WACV2021/html/Birhane_Large_Image_Datasets_A_Pyrrhic_Win_for_Computer_Vision_WACV_2021_paper.html. 2

[10] Rupert Brown, Susan Condor, Audrey Mathews, Gillian Wade, and Jennifer Williams. Explaining intergroup differentiation in an industrial organization. *Journal of Occupational Psychology*, 59(4):273–286, 1986. 3

[11] Amy Bruckman. Studying the amateur artist: A perspective on disguising data collected in human subjects research on the Internet. *Ethics and Information Technology*, 4(3):217–231, Sept. 2002. 3

[12] Joy Buolamwini and Timnit Gebru. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Conference on Fairness, Accountability and Transparency*, pages 77–91. PMLR, Jan. 2018. http://proceedings.mlr.press/v81/buolamwini18a.html. 1, 2

[13] Kathy Charmaz. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Pine Forge Press, Thousand Oaks, CA, Jan. 2006. 3

[14] Adrian Cho. United States establishes a dozen AI and quantum information science research centers. *Science — AAAS*, Aug. 2020. https://www.sciencemag.org/news/2020/08/united-states-establishes-dozen-ai-and-quantum-information-science-research-centers. 1

[15] Sasha Costanza-Chock. Design Justice, A.I., and Escape from the Matrix of Domination. *Journal of Design and Science*, July 2018. 2

[16] Kathleen M. DeWalt and Billie R. DeWalt. Participant Observation: A Guide for Fieldworkers. In *Participant Observation: A Guide for Fieldworkers*, pages 120–140. Altamira Press, Lanham, MD, USA, 2002. 2, 3

[17] Holly Else. AI conference widely known as 'NIPS' changes its controversial acronym. *Nature*, Nov. 2018. https://www.nature.com/articles/d41586-018-07476-w. 8

[18] Steven Feldstein. The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression. *Journal of Democracy*, 30(1):40–52, 2019. 1

[19] Pedro F. Felzenszwalb, Ross B. Girshick, David McAllester, and Deva Ramanan. Object Detection with Discriminatively Trained Part-Based Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1627–1645, Sept. 2010. 4

[20] Deborah B. Gould. *Moving Politics: Emotion and Act up's Fight against AIDS*. The University of Chicago Press, Chicago, IL, 2009. 2

[21] Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M. Branham. Gender Recognition or Gender Reductionism? The Social Implications of Embedded Gender Recognition Systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 1–13, New York, NY, USA, Apr. 2018. Association for Computing Machinery. 2

[22] Mark Harris. NTSB Investigation Into Deadly Uber Self-Driving Car Crash Reveals Lax Attitude Toward Safety - IEEE Spectrum. https://spectrum.ieee.org/cars-that-think/transportation/self-driving/ntsb-investigation-into-deadly-uber-selfdriving-car-crash-reveals-lax-attitude-toward-safety, Nov. 2019. (accessed 2021-03-18). 1

[23] Kenneth Holstein, Jennifer Wortman Vaughan, Hal Daumé, Miro Dudik, and Hanna Wallach. Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 1–16, New York, NY, USA, May 2019. Association for Computing Machinery. 2, 5

[24] Lilly Irani. *Chasing Innovation: Making Entrepreneurial Citizens in Modern India*. Princeton Studies in Culture and Technology. Princeton University Press, Princeton, NJ, 2019. 2

[25] Sean Kross and Philip J. Guo. Practitioners Teaching Data Science in Industry and Academia: Expectations, Workflows, and Challenges. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 1–14, New York, NY, USA, May 2019. Association for Computing Machinery. 2

[26] John Langford. The NIPS Experiment. https://cacm.acm.org/blogs/blog-cacm/181996-the-nips-experiment/fulltext, Jan. 2015. (accessed 2021-03-18). 1

[27] Zachary C. Lipton and Jacob Steinhardt. Research for Practice: Troubling Trends in Machine-Learning Scholarship. *Communications of the ACM*, 62(6):45–53, May 2019. 7

[28] Kristin Luker. *Salsa Dancing into the Social Sciences:*

Research in an Age of Info-Glut. Harvard University Press, Cambridge, 2009. http://qut.eblib.com.au/patron/FullRecord.aspx?p=3300119. 3

[29] Cheryl Mattingly. *Healing Dramas and Clinical Plots: The Narrative Structure of Experience*. Number 7 in Cambridge Studies in Medical Anthropology. Cambridge University Press, Cambridge, UK ; New York, NY, USA, 1998. 2

[30] Lokman I. Meho. E-Mail Interviewing in Qualitative Research: A Methodological Discussion. *Journal of the American Society for Information Science and Technology*, 57(10):1284–1295, 2006. 2

[31] Jacob Metcalf, Emanuel Moss, and danah boyd. Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics. *Social Research: An International Quarterly*, 86(2):449–476, 2019. https://muse.jhu.edu/article/732185. 2

[32] Cade Metz. A.I. Researchers Are Making More Than $1 Million, Even at a Nonprofit. *The New York Times*, Apr. 2018. https://www.nytimes.com/2018/04/19/technology/artificial-intelligence-salaries-openai.html. 1

[33] Cade Metz. When the A.I. Professor Leaves, Students Suffer, Study Says. *The New York Times*, Sept. 2019. https://www.nytimes.com/2019/09/06/technology/when-the-ai-professor-leaves-students-suffer-study-says.html. 1

[34] Rachel Metz. How one employee's exit shook Google and the AI industry. *CNN Business*, Mar. 2021. https://www.cnn.com/2021/03/11/tech/google-ai-ethics-future/index.html. 1

[35] Elliot G. Mishler. *Research Interviewing: Context and Narrative*. Harvard University, Press, Cambridge, Mass London, 1986. 2

[36] Michael Muller, Ingrid Lange, Dakuo Wang, David Piorkowski, Jason Tsay, Q. Vera Liao, Casey Dugan, and Thomas Erickson. How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 1–15, Glasgow, Scotland Uk, May 2019. Association for Computing Machinery. 2

[37] Sarah O'Meara. Will China lead the world in AI by 2030? *Nature*, (7770):427–428, Aug. 2019. https://www.nature.com/articles/d41586-019-02360-7. 1

[38] Devi Parikh. Humans of AI: Stories, Not Stats. https://www.cc.gatech.edu/~parikh/humanstoriesai/, 2021. (accessed 2021-03-18). 2

[39] Samir Passi and Steven J. Jackson. Trust in Data Science: Collaboration, Translation, and Accountability in Corporate Data Science Projects. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):136:1–136:28, Nov. 2018. 2

[40] Raymond Perrault, Yoav Shoham, Erik Brynjolfsson, Jack Clark, John Etchemendy, Barbara Grosz, Terah Lyons, James Manyika, Saurabh Mishra, and Juan Carlos Niebles. The AI Index 2019 Annual Report. Technical report, Human-Centered AI Institute, Stanford, CA, Dec. 2019. https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf. 1

[41] Charles C. Ragin. *Constructing Social Research: The Unity and Diversity of Method*. Sociology for a New Century. Pine Forge Press, Thousand Oaks, Calif, 1994. 2

[42] Morgan Klaus Scheuerman, Jacob M. Paul, and Jed R. Brubaker. How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):144:1–144:33, Nov. 2019. 2

[43] Sivic and Zisserman. Video Google: A Text Retrieval Approach to Object Matching in Videos. In *Proceedings Ninth IEEE International Conference on Computer Vision*, pages 1470–1477 vol.2, Oct. 2003. 4

[44] Victoria Song. MIT Takes Down Popular AI Dataset Due to Racist, Misogynistic Content. https://gizmodo.com/mit-takes-down-popular-ai-dataset-due-to-racist-misogy-1844244206, July 2020. (accessed 2021-03-18). 1

[45] Andrew Tomkins, Min Zhang, and William D. Heavlin. Reviewer bias in single- versus double-blind peer review. *Proceedings of the National Academy of Sciences*, 114(48):12708–12713, Nov. 2017. 2

[46] Antonio Torralba and Alexei A. Efros. Unbiased Look at Dataset Bias. In *CVPR 2011*, pages 1521–1528, June 2011. 2

[47] Antonio Torralba, Rob Fergus, and Bill Freeman. 80 Million Tiny Images. https://groups.csail.mit.edu/vision/TinyImages/, June 2020. (accessed 2021-03-18). 1

[48] Kiri Wagstaff. Machine Learning that Matters. *arXiv:1206.4656 [cs, stat]*, June 2012. http://arxiv.org/abs/1206.4656. 2

[49] Sarah Myers West, Meredith Whittaker, and Kate Crawford. Discriminating Systems: Gender, Race, and Power in AI. Technical report, AI Now Institute, 2019. https://ainowinstitute.org/discriminatingsystems.html. 1

[50] Li Yuan. How Cheap Labor Drives China's A.I. Ambitions. *The New York Times*, Nov. 2018. https://www.nytimes.com/2018/11/25/business/china-artificial-intelligence-labeling.html. 1

[51] Amy X. Zhang, Michael Muller, and Dakuo Wang. How do Data Science Workers Collaborate? Roles, Workflows, and Tools. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1):022:1–022:23, May 2020. 2