De-Mystifying Good Research and Good Papers (repost)

By Fei-Fei Li, Professor at Stanford, 2009-03-01

Please remember this:
1000+ computer vision papers get published every year!
Only 5-10 are worth reading and remembering!

Since many of you are writing your papers now, I thought that I'd share these thoughts with you. I probably have said all these at various points during our group and individual meetings. But as I continue my AC reviews these days (that's 70 papers and 200+ reviews — between me and my AC partner), these following points just keep coming up. *Not enough people conduct first class research. And not enough people write good papers*.

Every research project and every paper should be conducted and written with one singular purpose: . So when you conceptualize and carry out your work, you need to be constantly asking yourself this question in the most critical way you could – **"Would my work define or reshape xxx (problem, field, technique) in the future?"** This means publishing papers is NOT about "this has not been published or written before, let me do it", nor is it about "let me find an arcane little problem that can get me an easy poster". It's about **"if I do this, I could offer a better solution to this important problem," or "if I do this, I could add a genuinely new and important piece of knowledge to the field."** You should always conduct research with the goal that it could be directly used by many people (or industry). In other words, **your research topic should have many 'customers', and your solution would be the one they want to use**.

A good research project is **not about the past** (i.e. obtaining a higher performance than the previous N papers). **It's about the future** (i.e. inspiring N future papers to follow and cite you, N->inf).

A CVPR'09 submission with a Caltech101 performance of 95% received 444 (3 weakly rejects) this year, and will be rejected. This is by far the highest performance I've seen for Caltech101. So why is this paper rejected? Because **it doesn't teach us anything, and no one will likely be using it for anything**. It uses a known technique (at least for many people already) with super tweaked parameters custom-made for the dataset that is no longer a good reflection of real-world image data. It uses a BoW representation without object level understanding. All reviewers (from very different angles) asked the same question **"what do we learn from your method?"** And the only sensible answer I could come up with is that Caltech101 is no longer a good dataset.

Einstein used to say: **everything should be made as simple as possible, but not simpler**. Your method/algorithm should be the most simple, coherent and principled one you could think of for solving this problem. **Computer vision research, like many other areas of engineering and science research, is about problems, not equations**. No one appreciates a complicated graphical model with super fancy inference techniques that essentially achieves the same result as a simple SVM - unless it offers deeper understanding of your data that no other simpler methods could offer. A method in which you have to manually tune many parameters is not considered principled or coherent.

This might sound corny, but it is true. You're PhD students in one of the best universities in the world. This means you embody the highest level of intellectualism of humanity today. This means **you are NOT a technician and you are NOT a coding monkey**. When you write your paper, you communicate and . That's what a paper is about. This is how you should approach your writing. You need to feel proud of your paper not just for the day or week it is finished, but many for many years to come.

Set a high goal for yourself – **the truth is, you can achieve it as long as you put your heart in it!** When you think of your paper, ask yourself this question: Is this going to be among the 10 papers of 2009 that people will remember in computer vision? If not, why not? The truth is only 10+/-epsilon gets remembered every year. **Most of the papers are just meaningless publication games**. A long string of mediocre papers on your resume can at best get you a Google software engineer job (if at all – 2009.03 update: no, Google doesn't hire PhD for this anymore). **A couple of seminal papers can get you a faculty job in a top university**. This is the truth that most graduate students don't know, or don't have a chance to know.

Review process is highly random. But **there is one golden rule that withstands the test of time and randomness — badly written papers get bad reviews**. Period. It doesn't matter if the idea is good, result is good, citations are good. Not at all. **Writing is critical** — and this is ironic because engineers are the worst trained writers among all disciplines in a university. You need to discipline yourself: **leave time for writing, think deeply about writing, and write it over and over again till it's as polished as you can think of.**

Last but not the least, please remember this rule: **important problem (inspiring idea) + solid and novel theory + convincing and analytical experiments + good writing = seminal research + excellent paper**. If any of these ingredients is weak, your paper, hence reviewer scores, would suffer.