Digital Image Manipulation Forensic

Yan Zhao
Anthony Sutardja
Omar Ramadan

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2015-85
http://www.eecs.berkeley.edu/Pubs/TechRpts/2015/EECS-2015-85.html

May 13, 2015

DIMF

DIGITAL IMAGE MANIPULATION FORENSICS

UC BERKELEY MASTER OF ENGINEERING

EECS CAPSTONE PROJECT

Omar Ramadan

Anthony Sutardja

Yan Zhao

Advisor: Professor James F. O'Brien

Table of Contents

For the UC Berkeley Master of Engineering Capstone Project

Introduction

Problem Statement

For the past century, photographs have served as reliable primary sources of evidence, but that is quickly changing. Photo manipulation tools have become widespread and it is easy to manipulate images. Photo manipulations tools such as Adobe PhotoShop afford greater artistic expression, and enable users to create manipulations that challenge the limits of our natural perception. The difference between authentic and manipulated photos has become harder to distinguish, and can only be detected by digital forensic experts. The image forensics capstone project aims to create an online software service that performs the work of forensic analysts, and visualizes and analyzes the possible manipulations that may have been performed on an image. Our goal is to empower anybody to perform image forensic analysis, and help increase our faith in digital content.

Screenshots of our web detection service in action. *Left*: the landing page where a user can upload a suspected image. *Middle*: the progress page where users await detection results. *Right*: a section of the results page with our image manipulation analysis.

Our capstone has completed a prototype of this image manipulation detection web service that helps a user identify fraudulent features within an image. Although fully functional, our service is not yet ready for commercialization as many of the techniques are still under research. In the following sections, we perform a market, industry, and intellectual property analysis on how we would hypothetically take our online software service to market.

Trends, Market, and Strategy

Trends

There are strong technological and regulatory trends that reveal a growing need for effective image forensics solutions. Smartphones and the cameras embedded within them have rapidly become the most popular method of photo capture. According to the Pew Research Center, the percentage of U.S. adults who own smartphones jumped from 35% in 2011 to 56% in 2013 (Smith, "Smartphone Ownership 2013"). This relentless growth in smartphone adoption will continue, especially in regions outside of the United States (Dulaney 5). More people have the ability to take digital photos, as seventy-six percent of smartphone users say they use the smartphone's cameras to take photos (Smith, "Mobile Access 2010"). This instant and convenient photo-snapping ability is becoming recognized by many institutions as a way of documenting incidents and filing claims ("How to Document Auto Accident Damage.") (Thomson 3). However, as we increase our reliance on digital photography, we become susceptible to fraud. Digital image editing software is easy to obtain and is growing in utilization as well (Kahn 9). With the growing trend in smartphone adoption and the rise in access to image editing software, digitally manipulated images will become commonplace.

Due to the growing trend of digital image documentation, the regulatory landscape is also changing with new court precedents that are beginning to require verifying the authenticity of digital images. In Israel, a new law makes it illegal "to use images in advertisements that have been retouched to make models look thinner without printing a disclosure on the picture ("Picture Imperfect")". In the United States, the Federal Rules of Evidence now require the authentication of digital images before they are used as evidence in court (Thomson 3). The growing trend in regulatory measures for verifying digital images will only continue to push the need for image forensics service.

Hence, identifying these trends has helped us realize the customers of the image forensics industry. In journalism, we see the need to verify digital images for journalistic integrity. In the courtroom, we see the need to verify digital images to establish evidentiary authenticity. In insurance, we see the need to identify fraudulent digital insurance claims. As more services begin to rely on digital documentation, the need for such media verification services will increase.

Market

Potential Markets

There are several markets in which an image forensics solution is greatly needed.

Our technology has serious implications for the press image market. The exploding growth of the Internet and digital content has empowered citizen journalism and transformed the press industry with exclusive rich media collections. With this growth, press companies now have the additional challenge of maintaining their credibility by validating media from new unverified sources. Photo editors at such institutions work to ensure that only genuine photos are published, but with advanced editing tools, even an experienced analyst can fail to identify image manipulations. The $32 billion dollar newspaper industry takes photo manipulation as a serious threat because of the damage it brings to their credibility (McKenna). Their zero tolerance policy towards photo manipulation frequently results in hundreds of thousands of dollars in losses; each manipulated photo is issued a "mandatory kill" that alerts all affected customers to control damages (Lum). The involved photographers are usually terminated such as in the case of Getty freelancer Marc Feldman (Lum), L.A Times staff photographer Colin Crawford (Irby), or Reuters freelancer Adnan Hajj ("Altered Images Prompt Photographer's Firing").

Another potential market resides in the court of law. Admitting manipulated photographic evidence in court is both a growing trend and a growing concern. Perjury is not uncommon, and police officers are estimated to commit perjury twenty to fifty percent of the time on fourth amendment issues (Slobogin). Lawyers also commit perjury, and even good lawyers and are untruthful in their profession. As it stands, the authenticity of photographs is contingent on the honesty of the witnesses. An image forensics expert is only called upon when there are no witnesses to testify in a lawsuit or when the photographic evidence has its integrity challenged (Kashi). Otherwise, there is no standard procedure to have the photo evidence thoroughly vetted. The $400 billion dollar legal industry needs a better way to admit photographic evidence in a court of law that is more robust than a single testament of the truth (Novet). We see our image forensic software as an opportunity to change the way photo evidence is processed.

Segmented Market Entry

The last market that we have identified is with insurance companies. Our capstone team has decided to focus on catering our image forensics software service towards the insurance segment due to the great value we will provide to insurance companies. This value stems from the major losses insurance companies face due to fraudulent claims. While digitally documented photography and documentation has made it more cost effective for the insurance industry to survey damages, it has exposed them to significant losses. This shift towards digital evidence has made it easier to tamper with photographs to exaggerate or fabricate insurance claims. One study found that one in seven hundred general insurance claims has been digitally altered, and one in seventy-five property damage insurance claims has been digitally altered ("Picture Imperfect"). These digitally altered insurance claims account for a significant portion of the $40 billion dollars lost by the insurance industry every year in the United States (FBI). These monumental losses indicate that insurance companies have strong needs for innovative ways to combat fraudulent claims. Our image forensics software seeks to address this problem by helping insurance companies identify fraudulent claims accurately, and in a cost effective manner.

Marketing Strategy

Product

The detection of manipulated images is an important problem that requires the consultation of highly specialized forensic experts. Given the growing nature of this problem, we propose a service that puts the analytical power of image forensic methods in the hands of the common user. We generate detailed reports that can be interpreted by non-experts, effectively allowing lower level analysts to create more value while relieving the workloads of busy forensic experts.

Our product increases the capacity of analysts; it provides them with forensic information that helps them make informed decisions when evaluating claims. While images submitted with a claim are only a portion of the evidence that analysts use to base their decisions, the information we provide helps them identify and flag suspicious claims for further review, ultimately enabling them to prevent more fraud.

Value Guarantee

The largest risk for our customers is a negative return on investment, in that the cost of our platform is not exceeded by the value of fraud that we stop. To reduce this risk, and promote the sale of our service, our strategy is to minimize the software acquisition risks for our customers. We can reduce the switching costs by supplying our own forward deployed engineers to integrate with our customer's software stacks and our own support team to provide training services. Furthermore, with efficiency our service offers and its effectiveness in combating fraud, we can offer a guaranteed return on investment.

Customer Outreach

Customer acquisition is the largest challenge that we will face. Given the business-to-business nature of our business model and that customers are likely to have a large bureaucracy with several decision makers, our sales timeline will likely take months. To generate leads, we plan to attend digital forensic conferences and insurance fraud conferences to promote our services.

Industry

The image forensics services industry is sub-industry of the much larger digital fraud detection services industry. By focusing on the insurance market first, we balance our needs for large potential customer volume with our customers' urgent needs for a solution to their fraud problem. With a large number of insurance corporations across the country that face large insurance fraud losses, we can mitigate the bargaining power of buyers and justify big ticket sales. Furthermore, the alternatives to our service are weak in this market, so the barriers to entry are low. Additionally, our software as a service (SaaS) model allows us to replicate and scale our offerings with ease.

Substitute Offerings

The use of digital evidence has become accepted practice. In the case of automobile insurance claims, agencies actually encourage the behavior. The DMV.org procedures for filing an auto insurance claim suggests using a "smartphone or camera to take pictures at the scene" of the automobile damages ("How Car Insurance Companies Investigate Accident Claims"). These images are easily manipulated to exaggerate damages to increase liability or remove evidence to lessen liability. Insurance industry studies reveal that 10 percent or more of claims filed are fraudulent (GEICO, "GEICO's Special Investigations Unit"), many of which include digitally edited photos (Verifeyed, "Verifeyed – Image, Video and PDF Forensics, Authentication, Manipulation Detection, and and Digital Security"). By analyzing claims for fraudulent images, we can greatly improve the efficiency of the claim review process.

Current Methods

The details as to how insurance companies handle insurance claims varies based on the nature and severity of the claim, as well as the company's own policies. However, there are certain steps which are common amongst investigations.

Claims process begins when a report of damage is submitted to the insurance company using an online form or by directly calling the company's claim division. The claimant is expected to supply sufficient documentation that allows the company to understand the circumstances of the accident, and estimate the damages associated with it.

The case created is assigned to an employee of the insurance company. These employees have a number of titles across companies including Liability Investigators or Claim Adjusters, Appraisers, and Investigators (U.S. Bureau of Labor Statistics). However, the responsibilities of such personnel are fairly consistent; their role is to determine whether the insurance policy will pay out, and if so, how much it will cover.

The first step of evaluating the claim involves surveying the damage, and checking the case for inconsistencies. This is achieved through in-depth interviews with the involved parties and witnesses followed by a comparison of all accounts (GEICO, "How GEICO Investigates a Claim"). The interviews are also compared with information from the police report, and the investigator reviews auxiliary documents such as photos associated with the incident and public information from social media (GEICO). In the event that a case is flagged by an examiner as being suspicious, the case is reassigned to a Special Investigations Unit (SIU) which examines the case closely for evidence of fraud (GEICO). If the images associated with the claim come into question, the SIU will consult an image forensics expert.

The main problem in identifying fraudulent cases is the intense labor that it requires. Before a case is even flagged as being fraudulent and referred to the SIU, it has already consumed the resources of the claim adjuster. In the case of low-value claims where it is cheaper to pay the claimant rather than verify the damage, image verification can cut costs.

Manipulated images are dangerous to claim investigators. In the case of a fraudulent claim, it is unlikely that a perpetrator would submit photographic evidence that contradicts the claim story. Manipulated photographs can go unnoticed to the naked eye, and since "seeing is believing," convincing fakes can bring investigators to lapses in judgement. While there are forensic tools that can be used to verify photos, such tools require technical analysis skills that are beyond the scope of a case examiner's capabilities.

Another option insurance companies exercise in the face of fraud is to simply be passive. Losses are a common occurrence in the insurance market, and the way they are dealt with is by passing them onto their customers in the form of increased premiums (State Farm®). Not only does it make insurance less affordable, but according to the National Insurance Crime Bureau (NICB), insurance fraud is also correlated with higher taxes and inflated prices for consumer goods (National Insurance Crime Bureau).

Our solution in comparison

Our capstone solution provides an analytical report for photographs uploaded to our web service. Instead of using photographs as a means to validate a fraudster's narrative, we advocate using the images as one of the primary means for detecting fraud. As soon as a case is uploaded to the insurance company, the photographic contents can be automatically processed to allow case examiners and SIUs to quickly identify fraudulent claims from our reports. With our simple interface that localizes the manipulated regions within a photograph, we empower analysts to perform forensic tasks that were only accessible to experts. Our forensic service allow analysts to detect more fraud, and work through more cases by increasing their efficiency.

Our software-as-a-service platform can be integrated into the existing workflows of analysts, creating synergy. Compared to current substitutes for detecting fraud, the superior accuracy and labor savings that our workflow introduces is likely to yield a near immediate return on investment.

Competitive Landscape

Our capstone team is not the first group of people who have identified the value to be restored to the insurance companies from detecting manipulated images. We have identified two companies that our product will directly potentially compete with.

Competitors

System of Methods and Tools of Digital Processing Technology LLC, known as SMTDP is a Russian technology company founded in 2011 that focuses on automated business processes and image manipulation detections ("Company Overview of SMTDP Technology, LLC"). Some of the technologies used by SMTDP are image metadata analysis and image compression analysis. SMTDP's business mainly focuses on the development of technology and relies on partnership agreements with other companies, like Belkasoft and PricewaterhouseCoopers, to distribute and utilize their products ("Company Overview of SMTDP Technology, LLC"). Since SMTDP interacts only with value added resellers, the customers that receive the end user product depends entirely on these partners. Being based in Saint Petersburg, SMTDP is limited in its customer reach. SMTDP instead works with its partners to create third-party products that integrate their image manipulation detection methods

(SMTDP, "SMTDP | Company"). This is a model that leads to customer support and integration difficulties as opposed to working directly with customers to meet their needs.

Besides SMTDP, the second largest competitor in the digital image forensics space is Verifeyed, a privately owned company located at Czech Republic (Verifeyed, "Verifeyed | What Is Verifeyed"). Verifeyed's product is a software suite that can be installed by users who obtain a license; then end users can process suspect images to evaluate if the images are manipulated or not (Verifeyed, "Verifeyed | What Is Verifeyed"). Unlike SMTDP, Verifeyed distributes their software suite by themselves and sells directly to its customers. From the technology perspective, Verifeyed focuses on image metadata analysis and ballistics analysis (a type of image compression analysis).

Both SMTDP and Verifeyed have direct and indirect customers that are from major insurance companies, journalism industries, and credit card companies. The companies belong to markets that highly value image authenticity in order to reduce the potential losses to fraud. Like SMTDP and Verifeyed, our targeted markets are the insurance, journalism, and credit card markets, though we will target the insurance market segment first.

Comparison

Compared to our competitors' businesses and products, our image detection web service is strategically positioned to have more informative detection results and to be simpler in integration.

Our product is more informative in local manipulation changes than other products on the market. Our algorithms focus on both low level image analyses and high level image features, while our competitors mainly focus on bringing meaning out of a suspected image's metadata and compression analysis. Verifeyed's key differentiating factor is their image ballistics method which can identify the source camera of a photo using its quantization table. While this requires building a database of camera signatures, augmenting publicly available datasets is not an insurmountable challenge. We have matched some features of our competitors by providing the same low-level analysis, but we have also added more advanced forms of analysis. We look for high-level image features like Copy-Move detections and image-splicing detections. These high level features are common amongst most digital image manipulations, and are more difficult to conceal by image forgers. These extra dimensions help

our customers further in interpreting what happened to a particular image. The result of all these feature detections is a more comprehensive and informative image manipulation detection.

Our product has simple and interpretable results. Our competitors' software tools only outputs simple binary classification result indicating if an image has been manipulated or not. In contrast, our web service provides a detailed report showing visualizations of interesting manipulated features, as well as human-readable interpretations of what manipulations may have happened in the image. This detailed report is crucial to our customers since a simple but powerful tool would reduce the need to hire image analysis experts to investigate suspect images.

Screenshots of Verifeyed Professional Edition's user interface on Windows, which is not accessible through different platforms (Verifeyed).

Most importantly, our image detection software is a web service product that can be integrated into any existing workflow. SMTDP distributes its technology through third party resellers; its technology could only benefit users after integrating with third parties products (SMTDP, "SMTDP | Partners"). Verifeyed sells traditional softwares -- such that users need to purchase and install Verifeyed's software and install on their own computers (Verifeyed, "Verifeyed | What Is Verifeyed"). In contrast, our customers are not required to use a particular operating system to run a standalone application. Furthermore, the REST service's flexibility allows our product to be integrated into existing workflows helping to reduce customer training burdens and switching costs.

Combatting New Entrants

The largest threat that our product faces comes from academia. Academia is constantly working to develop new techniques that can offer incremental accuracy improvements. There is little we can do to mitigate this problem. That being said, we believe that the added utility of new accuracy improvements from another entrant would not be worth the switching costs to another platform.

The very nature of our detection service innately combats new entrants by using machine learning. Our system will become more accurate as we collect more data from our customers. Thus, the longer our system is in use, the better our accuracy becomes. Data acquisition of good quality for this space is not easy to obtain (as we have discovered from working on this capstone project). This first-mover advantage would make it difficult for new entrants to achieve our accuracy levels using the same techniques.

Scalability

Our software is designed to be deployed on a distributed infrastructure that can scale linearly with the rate at which we process images. We can grow and scale with the requirements of our customers.

Software-as-a-service (SaaS) products only have one supplier: cloud datacenters. The companies that offer cloud datacenters services have been relentlessly cutting prices to the point where all the providers are extremely cheap (State Farm®). If one provider tried to raise prices, we could easily shift our product offerings to a different provider since software service deployment architecture is generally the same. The cheap landscape of cloud service providers enables to host our detection service at minimal costs.

The alternative to deploying our software service on the cloud is to rent or buy our own servers. However, this would have to be at a price point in which the cloud offerings are no longer cheaper, which is possible depending on the usage of the servers (Leong 12). Even in the event where we must purchase servers, hosting and maintaining servers is a rather small fixed cost. Hence, the choice of hosting on cloud providers versus hosting on our own servers is not a decision that will constrain our profitability.

Intellectual Property

For this capstone, we have developed new algorithms for detecting and identifying manipulated features in a digital image. Our team has taken careful consideration of the intellectual property laws of the United States to protect the future prospects of our business. Copyright and trade secret laws in the United States can help us enforce our business model and keep our competitive advantage.

(No) Patent Strategy

Our team recognizes the difficulty in receiving a software patent claim. There are several criteria needed to acquire a patent, and the two criteria that challenge most software patents, including our own are nonobviousness and novelty. After analyzing each of these criterion, we believe that our web service, overarching detection process, and feature-specific sub-processes are not patentable.

Non-obviousness

US patent law requires a patent claim to be non-obvious "to a person having ordinary skill in the art" (35 USC 103). Software patents can be invalidated due to the ideas being of "common sense" or "obvious to try" as a next step (Perfect Web Technologies, Inc).

We hold concerns that our detection web service as a whole is "obvious to try." Although our detection service uses a combination of different methodologies, these techniques that we've created are heavily based on academic papers and resources that are in the public domain. Using all of these different features in a machine learning classifier could be seen as the next step for performing a general image manipulation detection classification. Hence, the United States Patent and Trademark Office may invalidate our claim to a patent due to nonobviousness.

Novelty

The other main challenge in obtaining software patents is establishing novelty. US patent laws requires that a patent claim to be filed before "the claimed invention was patented, described in a printed publication, or in public use, on sale, or otherwise available to the public" (35 USC 102).

Existing Patents

Although we believe our processes to be non-patentable, we must make sure that we are not infringing on any patent claims. We performed a thorough search of patents relevant to our image manipulation detection service. Although we were unable to find any patents that captured the essence of our end-to-end classification system, we were able to find a few more-specific patents that pertain to the features that our classification engine uses.

Looking into patents filed into other countries, we found that Chinese patent CN102609948A had many of the similar keywords and the same ultimate objective to one of our subprocesses related to Copy-Move detection (张华熊). The United States Patent and Trademark Office will look to foreign patents to determine if there is already prior art (United States. Dept. of Commerce). However, upon a closer reading of the technical specifications of the Chinese patent, we have determined that the Chinese patent would not be seen as prior art to our processes. Even though we are using some of the same image-level characteristics to deliver an analysis like Scale-Invariant Feature Transform (abbreviated as SIFT), our process for Copy-Move detection involves technical processes that are unrelated and different to the technical processes detailed in the Chinese patent. We believe this differentiation is enough such that this claim cannot be applied to our Copy-Move detection process in the United States courts, as extreme specificity results in narrow applicability of a patent's claims (LizardTech, Inc). Hence, we do not believe we are infringing on any prior claims. In addition to the differentiated procedures of the Chinese patent, we can also disregard this foreign patent claim since it is not enforceable in the United States.

We were also able to find existing patents in the United States such as US8200034B2 by New Jersey Institute of Technology, US7439989B2 by Microsoft, and US7720288B2 by Eastman Kodak that target double JPEG compression. What is unique to these methods is not the artifacts that they are detecting, but the classification methods being used. All of the patents listed describe unique processes for detecting double compression, while the method we employ uses distinct workflows for aligned and nonaligned JPEG (Bianchi). That being said, the specifics of how our technique detects JPEG double compression are distinct from the claims made in the patent. Hence, like in the Chinese patent case, we believe that our implementation does not infringe on the patents and their ways of extracting this feature.

Even though there are several relevant patents to how we extract various types of image manipulation features, we believe our techniques are differentiated enough such that we are not infringing on any patent claims.

Existing Publications

Having processes that do not infringe on existing patent claims does not mean that we necessarily have a novel and patentable claim. In addition to an existing patent search, the United States Patent and Trademark Office will also look to publications to determine the existence of prior art when determining whether or not to grant a patent (United States. Dept. of Commerce).

Many of our components are based on academic papers from external research groups that have heavily guided our implementations of detecting interesting image manipulation features. These features include higher order statistics, JPEG error level analysis, and double JPEG compression (Farid and Lyu) (Krawetz) (Bianchi and Piva). Although our copy-move detection technique was developed independently from knowledge that is documented by the academic publications of Irene Amerini at the University of Florence in Italy, the United States Patent and Trademark Office would still use these publications as evidence of prior art due to the substantial similarity of processes and its publication date being in 2011 (Amerini).

In addition to the prior art within the feature extraction context, prior art in publications also exists in the context of our full image detection service. A dissertation from Columbia University by Yu-Feng Hsu and Shih-Fu Chang captures the essence of using a variety of image features to perform a classification for image manipulation detection (Hsu and Chang 1). Although the dissertation use different features, Hsu and Chang describe a full system that is very similar to image detection service.

No Patents

The prior art found in existing publications further reduce our possible claims to a patent. Furthermore, our use of our classification system with these specific features can be seen as obvious to try. We conclude that both the overall classification system and the individual feature components for our classifier are to be non-patentable.

Trademark Law

We anticipate that other new entrants could mimic our image detection processes since the processes are not patentable. Our advantage over new entrants is the time that we have already spent in finding and tuning the optimal parameters. In order to minimize the threat of another entrant copying our systems, we will use trade secret protections to protect our "methods, techniques, processes, procedures, programs, or codes" (18 USC 1839). In other words, we will be able to protect the parameters in our machine learning process, the manipulated image datasets that test our methods, as well as the detailed lists on our customers. Protecting these assets of our business will be crucial to maintaining our competitive advantage.

Copyright Law

The last, and most important part, of our intellectual property strategy is to enforce the copyright protections offered by United States law.

Primarily, we will use copyright to enforce our business model of selling the license to view the image forensic reports produced by our web service. United States copyright law gives us the sole right to reproduce, distribute, and display our image analysis reports (17 USC 106). Successful enforcement of copyright protections would prevent the scenario in which a middleman takes advantage of our online software service by reselling or redistributing our reports indirectly to our potential customers. Rather than selling our software or our reports, we will sell the license to view the image analysis reports that our web service produces.

United States copyright will also protect our right to create any derivations of our product (17 USC 106). This would prevent any scenario in which a third party used our image analysis reports to make a slightly more comprehensive report or redressed the report by plagiarizing our metrics and analyses. Courts in the United States would enforce our copyright based on the substantially similarity of ideas that are contained in a potentially infringing work, as well as the "look and feel" of the potentially infringing work (Data East USA, Inc).

Hence, copyright is the strongest portion of our intellectual property strategies. The enforceability of copyright protections will help our business model flourish.

Overall Intellectual Property Strategy

All in all, intellectual property laws will be used by our capstone team to minimize threats from competing businesses and to enforce our business model. Although we are unlikely to receive any patent claims to our product, other intellectual property laws can help our business flourish. Trade secret protections will help keep our head-start in training and tuning our machine learning models. Lastly, copyright protections will help enforce that no third-party can redistribute or reproduce our image analysis reports that we will sell to our customers.

Works Cited

17 USC. Sec. 106. 2008. *Cornell University Law School*. Web. 26 Feb. 2015.

18 USC. Sec. 1839. 1996. *Cornell University Law School*. Web. 26 Feb. 2015.

35 USC. Sec. 102. 2012. *Cornell University Law School*. Web. 2 Mar. 2015.

35 USC. Sec. 103. 2011. *Cornell University Law School*. Web. 26 Feb. 2015.

"Altered Images Prompt Photographer's Firing." *Msnbc.com*. Accessed April 15, 2015. http://www.nbcnews.com/id/13165165/ns/world_news-mideast_n_africa/t/altered-images-prompt-photographers-firing/.

Bianchi, T., and A. Piva. "Image Forgery Localization via Block-Grained Analysis of JPEG Artifacts." *IEEE Transactions on Information Forensics & Security* 7, no. 3 (June 2012): 1003–17.

"Company Overview of SMTDP Technology, LLC." *bloomberg.com*. N.p., n.d. Web. 12 Apr. 2015.

Data East USA, Inc. v. Epyx, Inc. 862 F. 2d 204. Court of Appeals, 9th Circuit. 1988. *Google Scholar*. Web. 26 Feb. 2015.

Dong, Jing and Wei Wang. "CASIA Tampered Image Detection Evaluation Database." CASIA Tampered Image Detection Evaluation Database. CASIA. 16 Mar. 2015.

Dulaney, Ken, et al. "Predicts 2015: Mobile and Wireless" Gartner, Inc. Stamford, Connecticut. November 5, 2014. Print.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay. Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825-2830 (2011)

Farid, Hany, and Siwei Lyu. “Higher-Order Wavelet Statistics and Their Application to Digital Forensics.” Accessed December 19, 2014.

FBI, “Insurance Fraud.” *FBI*. N.p., n.d. Web. 12 Apr. 2015.

GEICO, “GEICO’s Special Investigations Unit.” *geico.com*. N.p., n.d. Web. 12 Apr. 2015.

GEICO, “How GEICO Investigates a Claim.” *geico.com*. N.p., n.d. Web. 12 Apr. 2015.

GEICO, “Join GEICO in the Fight against Insurance Fraud by Reporting Suspicious Activity.” *geico.com*. N.p., n.d. Web. 12 Apr. 2015.

“How Car Insurance Companies Investigate Accident Claims.” *DMV.org*. N.p., n.d. Web. 12 Apr. 2015.

“How to Document Auto Accident Damage.” *DMV.org*. N.p., n.d. Web. 12 Apr. 2015.

Hsu, Yu-Feng and Chang, Shih-Fu, “Detecting Image Splicing Using Geometry Invariants And Camera Characteristics Consistency” International Conference on Multimedia and Expo (ICME), Toronto, Canada, July 2006.

Hsu, Yu-Feng, and Chang, Shih-Fu, “Image Tampering Detection For Forensics Applications.” Columbia University, 2009. Print.

I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, G. Serra. A SIFT-Based Forensic Method for Copy–Move Attack Detection and Transformation Recovery. Information Forensics and Security, IEEE Transactions. Volume 6, Issue 3. IEEE. 17 March 2011.

Irby, Kenneth. “L.A. Times Photographer Fired Over Altered Image,” April 2, 2003.

Kahn, Sarah. “Design, Editing & Rendering Software Publishing in the US” IBISWorld. September 2014. Print. Accessed on February 17, 2015.

Kashi, Joe. “Authenticating Digital Photographs as Evidence: A Practice Approach Using JPEG Metadata.” *americanbar.org*. N.p., n.d. Web. 12 Apr. 2015.

Krawetz, Neal. “A Picture’s Worth… Digital Image Analysis and Forensics”. Hacker Factor Solutions. Black Hat Briefings USA 2007.

Leong, Lydia. “Technology Overview for Cloud Infrastructure as a Service” Gartner, Inc. Stamford, Connecticut. June 30, 2014. Print.

LizardTech, Inc. v. Earth Resource Mapping, Inc. 424 F. 3d 1336. Court of Appeals, Federal Circuit. 2005. *Google Scholar*. Web. 26 Feb. 2015.

Lum, Jessica. “Getty Photographer Terminated Over Altered Golf Photo.” *PetaPixel*. N.p., n.d. Web. 12 Apr. 2015.

McKenna, Farrell. *Newspaper Publishing in the US*. Melbourne, Australia: IBISWorld Services, 2014. *IBISWorld*. Web. 12 Apr. 2015.

National Insurance Crime Bureau, “INSURANCE FRAUD: UNDERSTANDING THE BASICS.” *National Insurance Crime Bureau* n. pag. Web. 12 Apr. 2015.

Ng, Tian-Tsong and Chang, Shih-Fu, “A model for image splicing,” In: IEEE International Conference on Image Processing, Singapore, 2004.

Novet, Jordan. “All Rise: The Era of Legal Startups Is Now in Session | VentureBeat | Entrepreneur | by Jordan Novet.” *venturebeat.com*. N.p., n.d. Web. 12 Apr. 2015.

Perfect Web Technologies, Inc. v. InfoUSA, Inc. 587 F. 3d 1324. Court of Appeals, Federal Circuit. 2009. *Google Scholar*. Web. 26 Feb. 2015.

“Picture Imperfect.” *The Economist* 9 Mar. 2013. *The Economist*. Web. 12 Apr. 2015.

Porter, Michael E. “The Five Competitive Forces that Shape Strategy” Harvard Business School Publishing Corporation. 2008. Print.

State Farm®, “Reporting Fraud – State Farm®.” *State Farm*. N.p., n.d. Web. 12 Apr. 2015.

“Silver Lining.” *The Economist*. *The Economist*. Web. 12 Apr. 2015.

Slobogin, Christopher. “TESTILYING: POLICE PERJURY AND WHAT TO DO ABOUT IT.” *University of Colorado Law Review, Inc.* 67.1037 (1996): n. pag. Print.

Smith, Aaron. “Mobile Access 2010.” *Pew Research Center’s Internet & American Life Project*. N.p., n.d. Web. 12 Apr. 2015.

Smith, Aaron. “Smartphone Ownership 2013.” *Pew Research Center’s Internet & American Life Project*. N.p., n.d. Web. 12 Apr. 2015.

Smith, Steven W. “Data Compression”. “The Scientist and Engineer's Guide to Digital Signal Processing”. San Diego, Calif.: California Technical Pub., 1997. Print.

SMTDP, “SMTDP | Company.” N.p., n.d. Web. 12 Apr. 2015.

SMTDP, “SMTDP | Partners.” N.p., n.d. Web. 12 Apr. 2015.

Tachibanaya, TsuruZoh. “Description of Exif File Format.” Exif File Format. 19 Dec. 1999. Web. 16 Mar. 2015.

Thomson, Lucy L. “Mobile Devices: New Challenges for Admissibility of Electronic Evidence” The SciTech Lawyer, Volume 9, Number 3, Winter/Spring 2013. Print.

U.S. Bureau of Labor Statistics, “Claims Adjusters, Appraisers, Examiners, and Investigators : Occupational Outlook Handbook: : U.S. Bureau of Labor Statistics.” N.p., n.d. Web. 12 Apr. 2015.

United States. Dept. of Commerce. The United States Patent and Trademark Office. “Prior Art”. Chapter 0900. Section 901. *The United States Patent and Trademark Office.* Web. 26 Feb. 2015.

V. Christlein, C. Riess, J. Jordan, C. Riess, E. Angelopoulou: “An Evaluation of Popular Copy-Move Forgery Detection Approaches”, IEEE Transactions on Information Forensics and Security, vol. 7, no. 6, pp. 1841-1854, 2012.

Verifeyed, “Verifeyed – Image, Video and PDF Forensics, Authentication, Manipulation Detection, and and Digital Security.” N.p., n.d. Web. 12 Apr. 2015.

Verifeyed, “Verifeyed – Image, Video and PDF Forensics, Authentication, Manipulation Detection, and and Digital Security | Technology & Documentation.” N.p., n.d. Web. 12 Apr. 2015.

Verifeyed, “Verifeyed – Image, Video and PDF Forensics, Authentication, Manipulation Detection, and and Digital Security | What Is Verifeyed.” N.p., n.d. Web. 12 Apr. 2015.

Ye, Shuiming and Sun Qibin, “Detecting Digital Image Forgeries By Measuring Inconsistencies of Blocking Artifact”, Institute for Infocomm Research, Singapore 119613.

张华熊, 胡洁, 薛福冰, 黄海. “Manipulation detection method for copy-paste distorted photo digital photos.” Patent CN 102609948 A. 16 Apr. 2014.

Appendix

Figure 1: The landing page for our prototype image detection service.

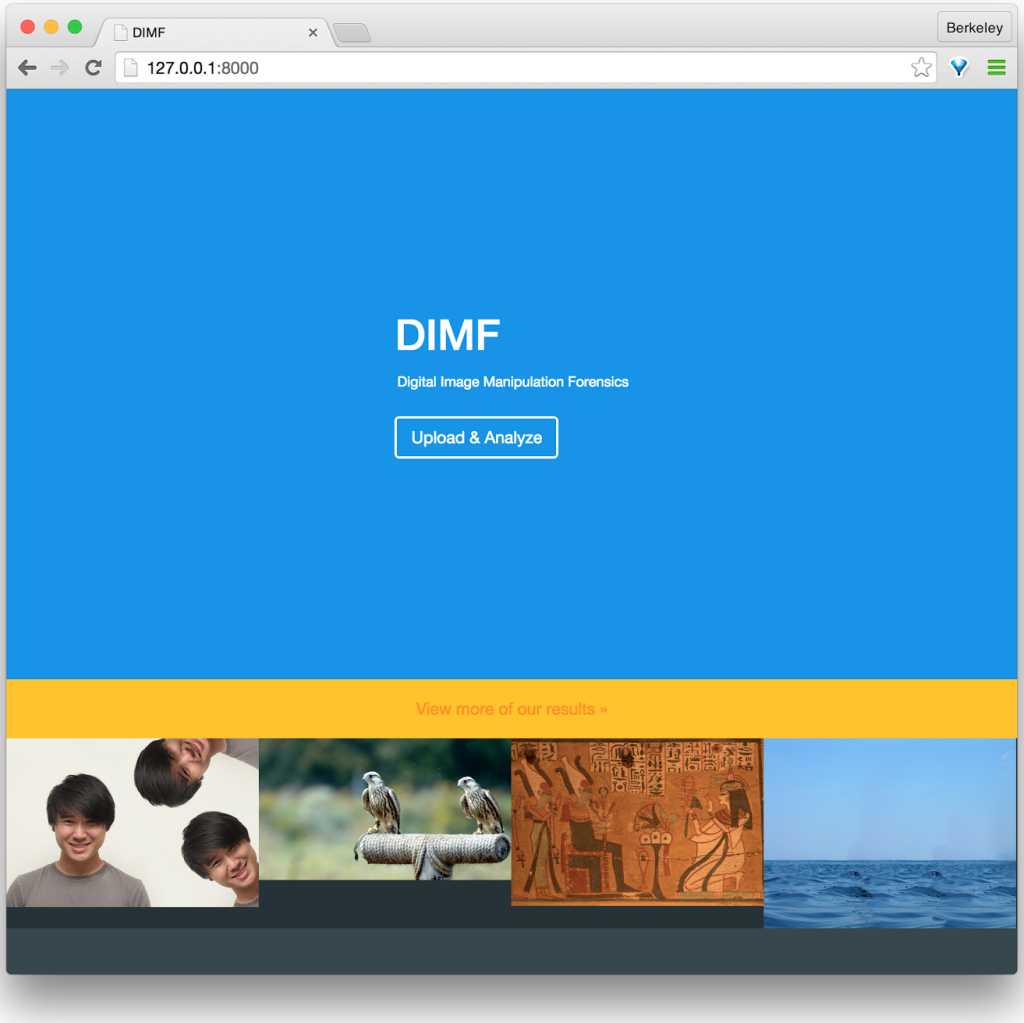


Figure 2: The progress page while a photo is being processed.

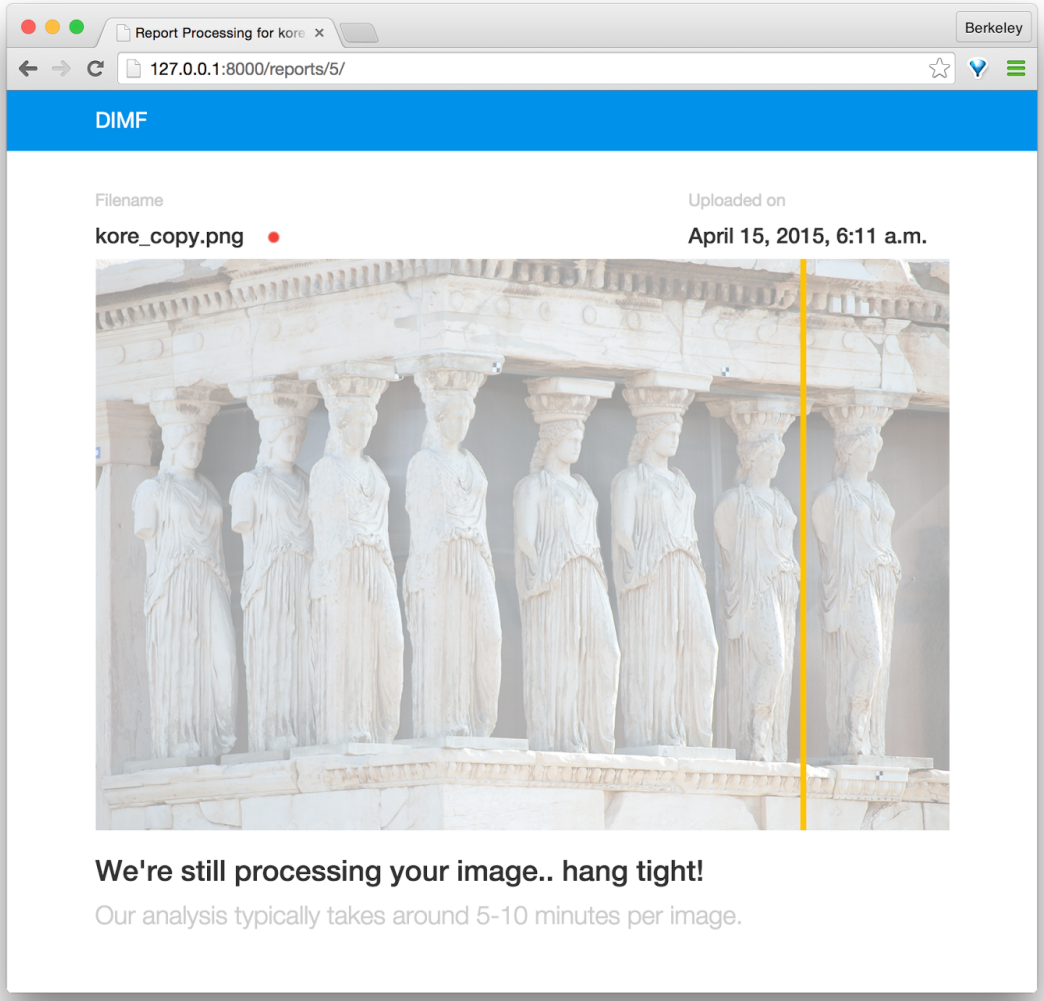
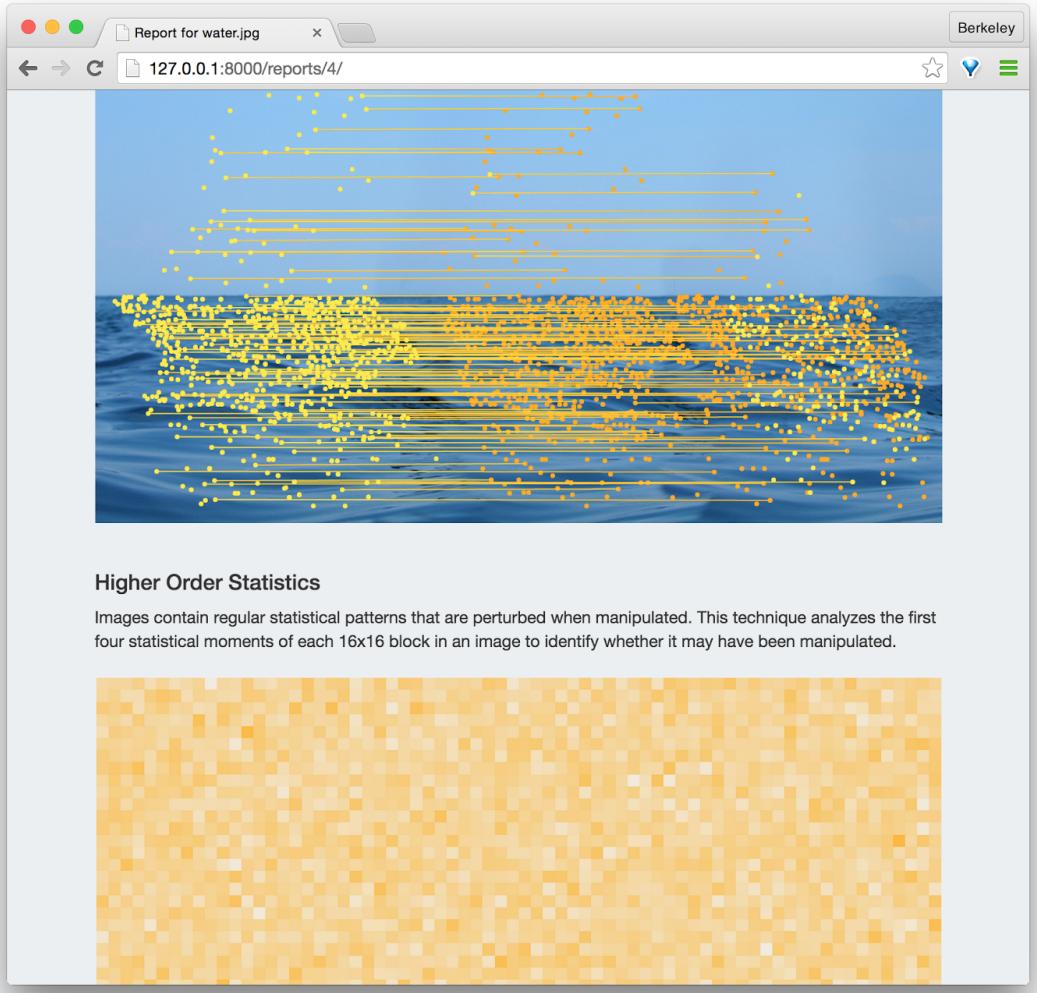


Figure 3: The results page detailing a copy-move detection and higher order statistic analysis.



Individual Technical Contribution

By Yan Zhao

Introduction

Digital images can be easily manipulated nowadays with sophisticated image manipulation software like Adobe Photoshop. On the other hand, image manipulation detection techniques have not improved a lot. This asymmetricity severely reduces the credibility of images as source of truth in many industries. Our capstone project is aiming to explore existing and develop new image manipulation detection techniques; furthermore, we would like to provide an easy-to-use software as a service product to average users to detect image manipulations.

As a team, we have found several aspects we can approach to detect image manipulations. Metadata analysis can be performed easily to detect traces left by photo manipulation softwares. At compression level, many image manipulation techniques will leave encoding artifacts that can be used to detect forgery activities. On image features level, there are many pixel based forgery detections for copy-move and image splicing, which are both used in many image manipulation cases.

Even though the team worked on many of the above aspects together, my technical contribution is mainly targeting on compression level analysis and web architecture of our software as a service product.

Image compression analysis

JPEG image format, the most common image format used today, uses lossy image compression; thus, every time a JPEG image is manipulated and resaved, some of its information will be lost due to image compression. In fact, as figure 1 shows, the JPEG image compression process consists of three steps that are discrete cosine transform, quantization, and entropy encoding (Smith). During the quantization step, many encoding artifacts will be left behind; more specifically, many image manipulation techniques will leave certain characteristic artifacts that can be used to prove whether an image has been manipulated or not.

Figure 1a. Image Encoding Process

Figure 1b. Image Decoding Process

Focusing on the quantization step, I have explored on two compression analysis techniques. The first one is to analyze blocking artifacts left during compression; and the second one is to perform error level analysis to find the inconsistency among different regions of a given image after re-saving.

Blocking artifact measurement

Literature Review

Blocking artifact is a kind of compression loss created during image compression. The visual appearance of a blocking artifact resembles some pixel blocks with a single color. Ye has claimed that image compression could serve as a "natural authentication code" to detect image manipulations (Ye). As figure 2 has shown, blocking artifact measurement is computed by taking the average of blocking artifact values over each 8 by 8 sub-blocks across the whole image. For each sub-block, difference between the actual DCT value and a rounded DCT value calculated from the quantization table will be measured.

$$B(i) = \sum_{k=1}^{64} |D(k) - Q(k)round(\frac{D(k)}{Q(k)})| \qquad BAM = \frac{1}{N}\sum_{i} B(i)$$

Figure 2. Equations used to compute blocking artifact of each individual block and the overall blocking artifact. The key point is to calculate the difference between the rounded DCT value calculated from quantization table and the actual DCT value calculated from transformation. After blocking artifact of each sub-block is computed, BAM can be calculated by taking the average value of the whole image (Ye).

Methodology

My first attempt is to reproduce Ye's result of blocking artifact measurement calculation. One obstacle I have encountered is retrieving the quantization table used for the given image, and it turns out the given JPEG image has quantization table provided as a metadata tag.

Since the blocking artifact measurement is more like an average of the blocking artifact values across the whole image. Visualizing blocking artifact values serves as a more helpful visual indicator than blocking artifact measurement, the numerical metrics.

Thus, in addition to extracting the blocking artifact measurement, I have visualized the blocking artifacts across the whole image. This process is performed by plotting each 8 by 8 sub-blocks with its blocking artifact value; thus, the blocking artifact image will be presented as a grayscale image indicating the blocking artifact at each sub-block. Figure 3 shows an example of blocking artifact plot on a manipulated image. In this image, the parachute is spliced into the desert image.

Figure 3. The image on the left is the given image, and the blocking artifact image is on the right. We can find the parachute region has higher blocking artifact values than that of the other parts of the images.

Results

Although the blocking artifact measurement provides some clues about whether an image has been manipulated or not, the technique is still relatively crude. If the given image has a low quality, the blocking artifact would be very high even for the unmanipulated parts. Also, during image

re-compression, the high frequency region, even unmanipulated, will still have a high blocking artifact because JPEG encoding will result higher losses in high frequency region.

Even though blocking artifact measurement cannot give a high accuracy result, it is a good indication on whether an image has been resaved or not because compressed image is usually re-saved in a lower quality.

Error Level Analysis

Literature Review

Because JPEG image encoding uses lossy compression, a certain amount of error will be introduced each time a JPEG image is compressed/resaved (Krawetz). Thus, each pixel resides in a certain error level, and performing a compression will change the error level of a given pixel. However, if a image is manipulated, it is highly likely that different parts in the image will have different error levels, and error level analysis is focusing on this artifact to identify manipulated image and find tampered region (Krawetz). Also, Krawetz has found that error level does not change linearly each time a JPEG image is saved -- such that saving an image twice with 80% loss is equivalent to saving the same image once with 64% of error (Krawetz). Therefore, error level analysis will resave the given image at a known error rate and calculate the difference between the given image and the resaved image. In many case, the manipulated region will have a different error level than the unmanipulated regions, and error level analysis will expose the manipulated region by labeling the regions with higher change in error level after resaving (Krawetz).

Krawetz's research on image manipulation detection is important. In his paper, Krawetz has first introduced error level analysis and pointed out several crucial observations about manipulated images and practical detection techniques.

Dataset

There are three image manipulation datasets I have used for error level analysis. The first one is "CASIA2 - Tampered Image Detection Evaluation Database"; CASIA2 dataset consists of 7491 authentic and 5123 tampered color images with various image manipulation techniques (Dong). The second one is "Columbia Uncompressed Image Splicing Detection Evaluation Dataset." that has 180

spliced images and 180 authentic images (Hsu). The third one is “Benchmark Dataset” by Christlein and her colleagues; “Benchmark Dataset” has very high quality copy-moved images (Christlein).

Work Reproduction

Figure 4 shows an example result of error level analysis I have generated with a manipulated image made by Anthony, my partner. I was able to reproduce Krawetz’s work with several manipulated and unmanipulated images. The analysis below shows my observations about and work beyond Krawetz’s proposals.

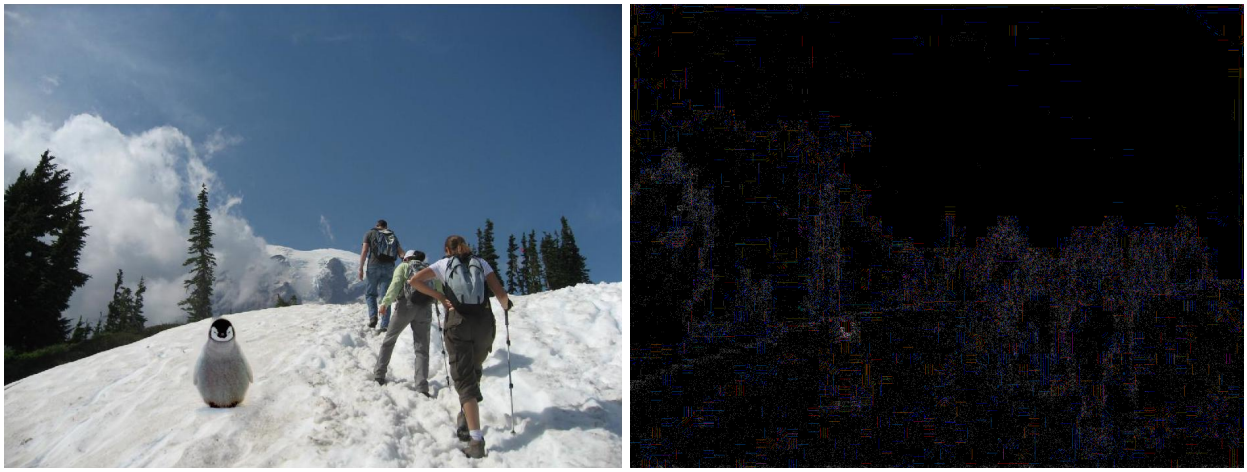


Figure 4. The image on the left is the given image, and an error level analysis (ELA) image is on the right. In the ELA image, regions with higher changes in error level have brighter pixels. Clearly, we can find the penguin’s nose is a suspicious region. Actually, the penguin is spliced into the given image from another image with poisson blurring.

Error level analysis works best for spliced images -- such that the manipulated part is spliced from other images because the spliced parts and original parts are highly likely to have different error levels (Ng). As for copy-move, another common image manipulation technique that copies part of an image and moves to other regions of the same image, error level analysis might not work very well since both of the tampered part and original part are from the same image. However, because image manipulation softwares will usually combine different layer together to generate the manipulated image, error level analysis could still find some traces left by copy move manipulations, as shown in figure 5.

Figure 5. The image on the left is the given image, and the error level analysis (ELA) image is on the right. In the given image, the bird on the cattle's head is copied from the bird on the ground. The ELA image shows a clear high region on the copied bird area.

Improvement & Result

Error level analysis does have a drawback - such that the high frequency region in the given image will affect the result. JPEG image encoding will naturally compress more in the high frequency regions of an image (Smith). Thus, the high frequency part of an image will usually have a relatively greater drop in error levels even though it is not tampered.

Low Frequency Mask

In order to overcome this drawback, a low frequency mask is generated by running gaussian filter on the image and measuring the difference between the given image and filtered image because the low frequency regions will have smaller changes after gaussian filter. Then this mask is multiplied on each of the three different channels of the error level analysis data. The purpose of this low frequency mask is to make the error levels analysis data of the low frequency regions to contribute more for visual analysis. Also, low frequency regions will be highlighted to present an alert of high risks. Figure 6 shows a visualization of a masked error level analysis image.

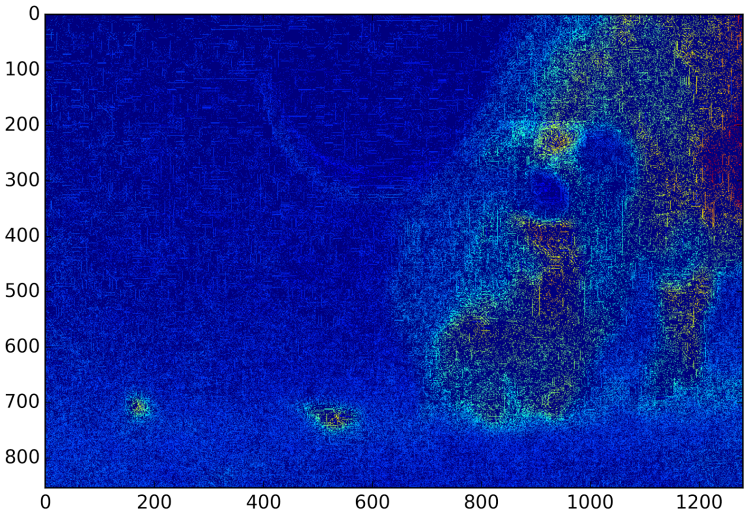


Figure 6. This heat map shows the masked ELA image. The original image is the cattle image in Figure 5. Clearly, we could find the low frequency region has a higher value in the heat map. More specifically, the manipulated bird on the cattle's head is showing a much brighter color.

In addition to weight ELA image with the frequency of the given image, another approach I have implemented is to label the low frequency region and the higher ELA region in the given image and identify the intersection. This approach is generally more effective because it is more visually friendly and does not ignore manipulated region with high frequency.

High Error Level Analysis Data Mask

Instead of presenting the raw error level analysis image, as shown in figure 5, a high error level analysis data mask is generated to better demonstrate the analysis results. First, the same low frequency mask will be generated as discussed above. Then, the low frequency mask will be gaussian filtered and thresholded to label low frequency regions.

Figure 7. Process of generating continuous high ELA region from ELA image data. L2-Norm will reduce the three channels ELA image to one channel. First thresholding will filter out low value ELA noise. Gaussian filter will then produce a continuous ELA region from scattered points. Finally, scaling up to 255 and thresholding again will further reduce noise.

In order to find a continuous high ELA region, I have first calculated the L2 Norm of the ELA image data as the magnitude. Then, the ELA magnitude data is thresholded to filter out noise values. After thresholding, ELA magnitude data is gaussian filtered to produce a continuous region. Finally, this smoothed magnitude data is thresholded again to remove noise. Figure 8 below shows the low frequency mask and high ELA mask on the given image. As we can see, labeling low frequency region and high ELA region is more visually friendly and is better at identifying suspicious region, which is the red region above the cattle's head in the example at figure 8.

Figure 8. The blue region shows the low frequency mask, and the yellow region shows the high ELA mask. The intersection between these two regions is labeled red. The red region has a higher probability to get manipulated.

Classification

Besides visual output, I have also extracted numerical metrics from error level analysis data for manipulated vs. authentic image classifier. I have used support vector machine and adaptive boosting classifier provided in "Scikit-learn" machine learning library (Pedregosa). After generating the ELA image, I have calculated the mean, median, and variance on each of the three channels in the ELA image. The histogram of the extracted metrics on CASIA2 dataset shows a clear separation between authentic and tampered images, as shown in figure 9 below.

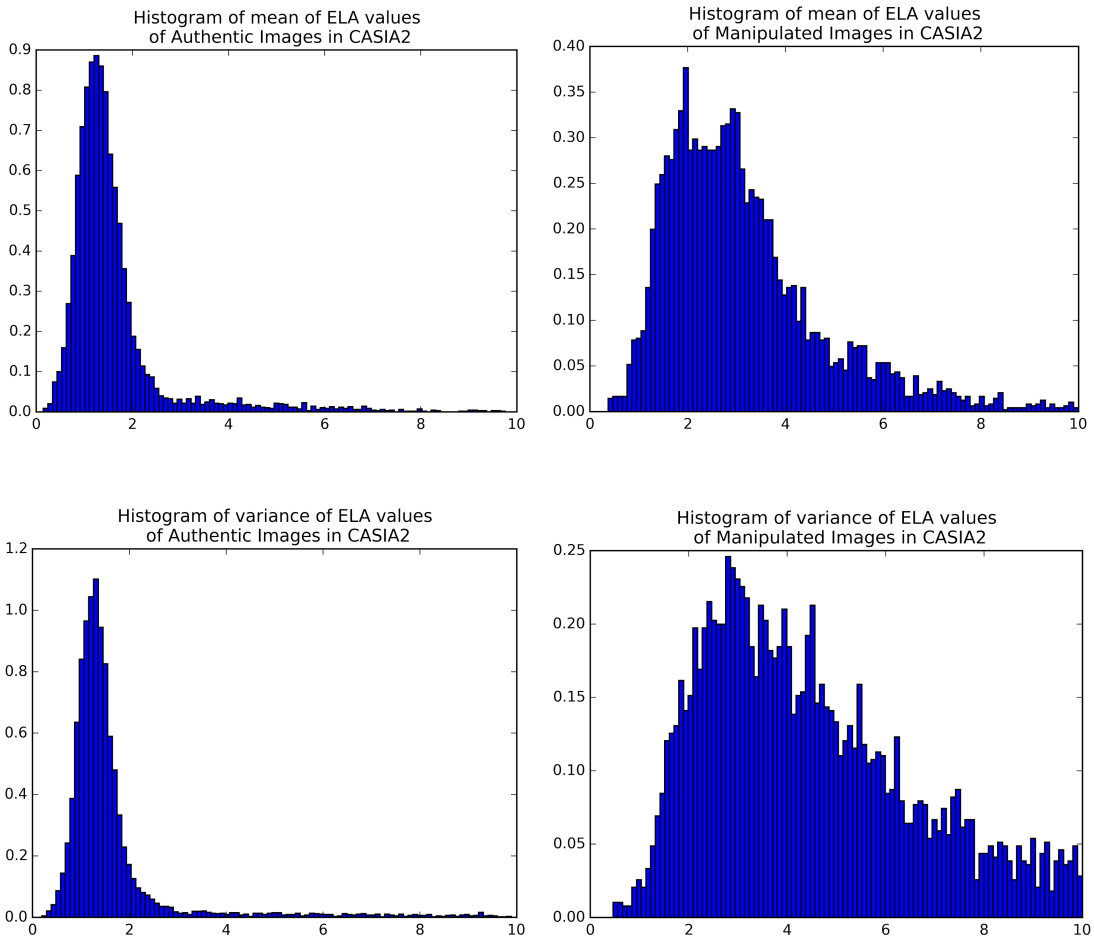


Figure 9. Histogram of different numeric metrics of ELA on CASIA2 dataset shows a clear separation between the manipulated and original images. The left two figures show histograms of mean and variance of original images. The right two figures show histograms of mean and variance of manipulated images.

With the extracted numeric metrics as sample feature, I have used support vector machine and adaptive boosting to classify test images (from CASIA2 dataset) to be either manipulated or authentic. With cross validation, support vector machine has a success rate of 84.34%, and adaptive boosting with decision tree has a success rate of 85.54%. Figure 10 below shows a plot of the decision output of both labeled classes. Unfortunately, the image classifier does not work well on the benchmark dataset generated by Christlein.

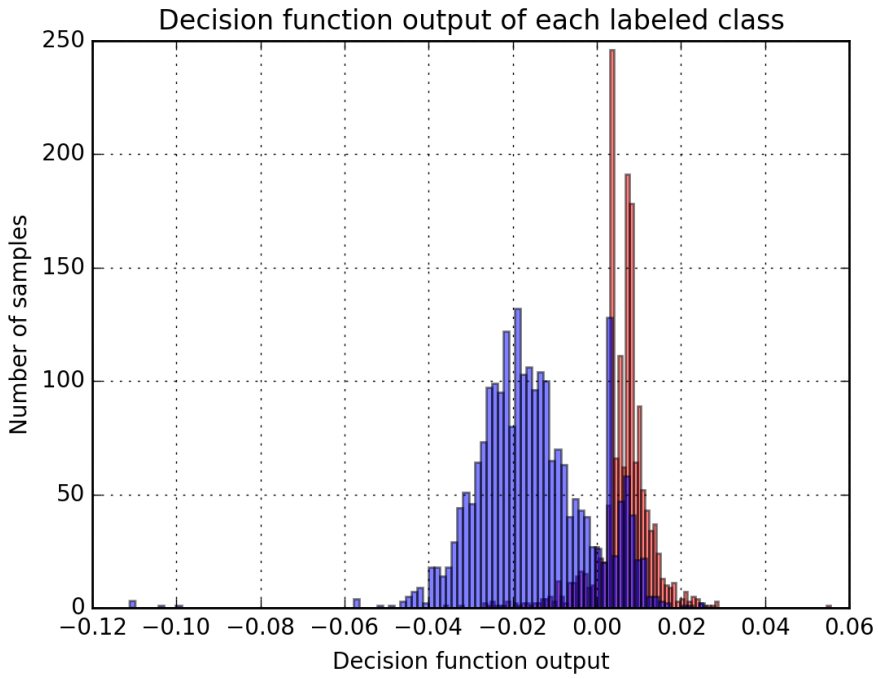


Figure 10. Histograms of decision function output using adaptive boosting of manipulated labeled image (red) and authentic labeled image (blue). Value 0 is the decision boundary such that images with output value below zero will be labeled as authentic and images with output value above zero will be labeled as manipulated. Even though there are some errors of falsely labeling authentic images as manipulated images, the classifier does perform a good job in separating the manipulated and labeled sample images.

Metadata Analysis

JPEG format not only stores the digital image, but also records various meta-data information along with the digital image file (Tachibanaya). Those metadata information includes resolution, creation time, modification time, software, and many other information.

Together with Omar and Anthony, my partners, we have performed metadata analysis by extracting metadata from given JPEG image, and we do find common fields we can use to identify if an image is manipulated or not.

Methodology & Result

Usually, the software section in metadata will record the software used to generate the digital image. In the case of smartphone, the operating system version will be stored; however, in the case of

manipulated images, the software, like Adobe Photoshop, used to perform the manipulations will be record as well.

Another important metadata information is the time of creation and modification. Even though some photo manipulation software could change the above timestamp fields, it is still likely that the given image is manipulated if these two fields do not contain the same timestamp.

However, one drawback of metadata is that the forgers could also change the metadata. Thus, some "authentic looking" metadata will not serve as a clean pass for an image; however, suspicious metadata will definitely be a strong indication of image manipulation.

Web Architecture

Besides performing compression based analysis on images, I have also focused on the design and implementation of the web service for our software as a service product.

Our web system contains two main parts. The first part is the web server that accepts images from users and triggers analysis worker. The second part is the analysis worker that performs image manipulation detection and generates reports asynchronous.

We have used python and matlab as our backend programming language for our first prototype.

Summary

As a team, we are trying to use machine learning techniques to provide an accurate and scalable image manipulation service. My work mainly focuses on image compression analysis, including blocking artifact measurement and error level analysis and the design and implementation of our web service.

Concluding Reflections

By Yan Zhao

Outcome

Image manipulation detection is a subtle work; like its opponent, image manipulation, detecting image manipulation is more like a combination of analysis and art work.

Our current output falls in line with what we originally planned. We have explored on various aspects of image manipulation detection techniques and extended many existing work. We have the beta version of our web service built up and could successfully generate image analysis report as we planned.

Future work

Future work on error level analysis consists of two parts. The first part is to build other tools to help users interpret error level analysis images. Currently, a low frequency mask and high error level analysis mask are generated to label the suspect regions. The other part is to extracting more meaningful numeric metrics from error level analysis images.