Forensic Methods for Detecting Image Manipulation - Copy Move

Anthony Sutardja
Omar Ramadan
Yan Zhao

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2015-84
http://www.eecs.berkeley.edu/Pubs/TechRpts/2015/EECS-2015-84.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.

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

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 Shih-Fu Chang. "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.

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.

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

SMTDP, "SMTDP | Partners." N.p., n.d. Web. 12 Apr. 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.

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.

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

Code

- Web service: https://github.com/ucb-image-forensics/Detection-Service
- Detection algorithms: https://github.com/ucb-image-forensics/imforensics

Figures

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

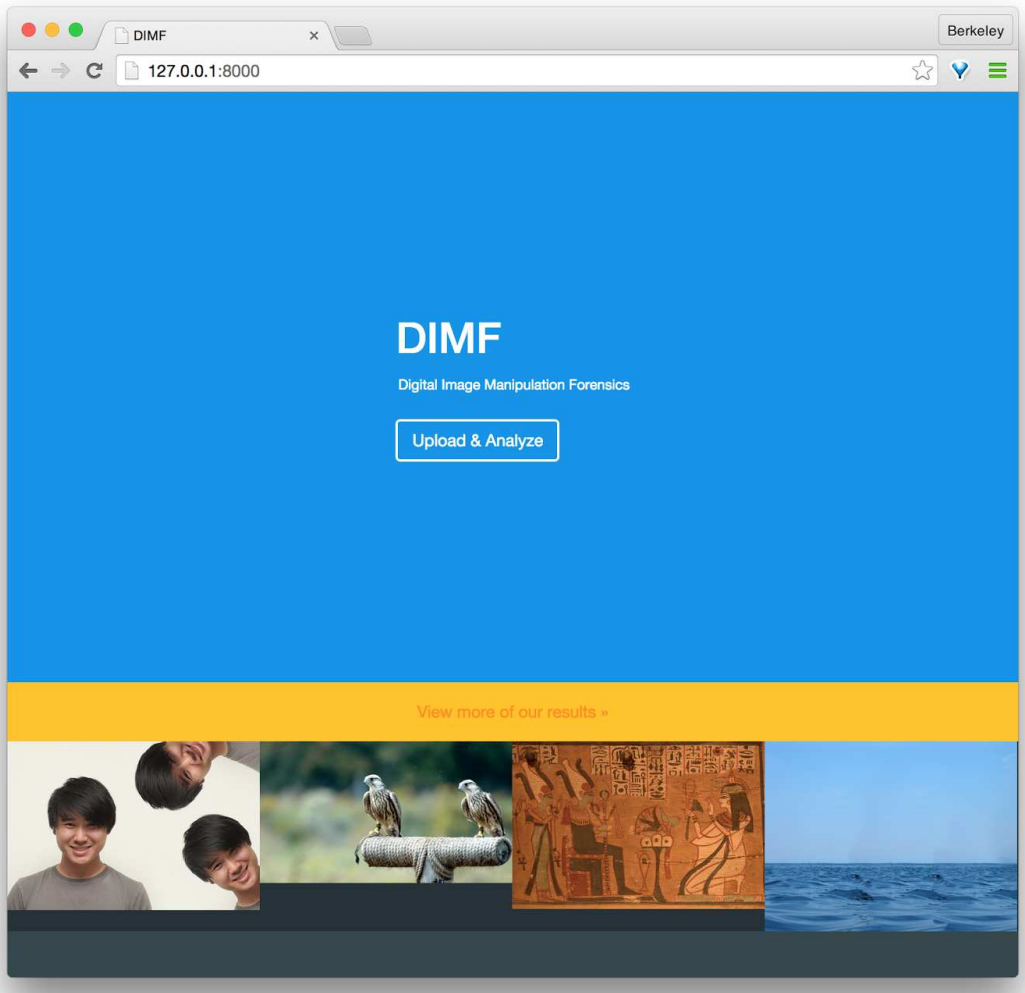


Figure 2: The progress page while a photo is being processed.

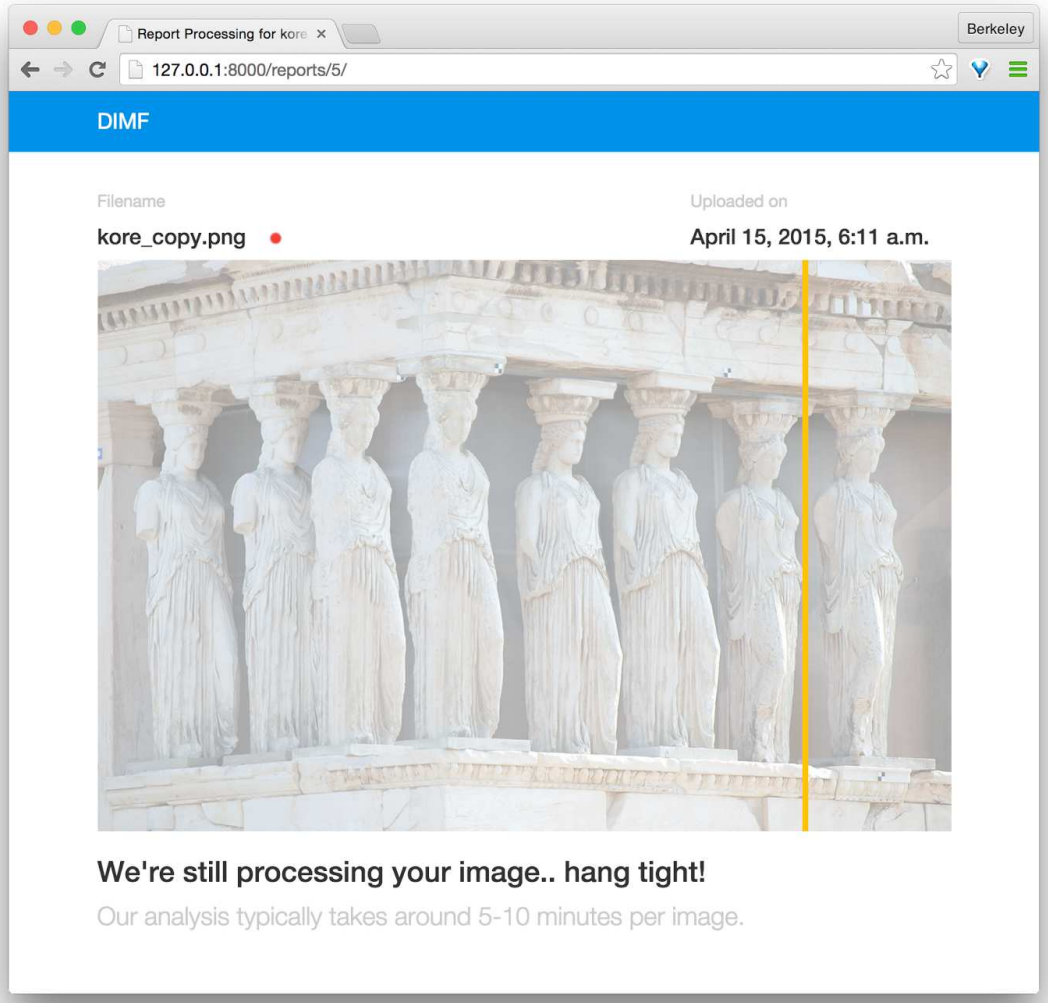


Figure 4: The results page detailing an overall analysis as well as the metadata analysis.

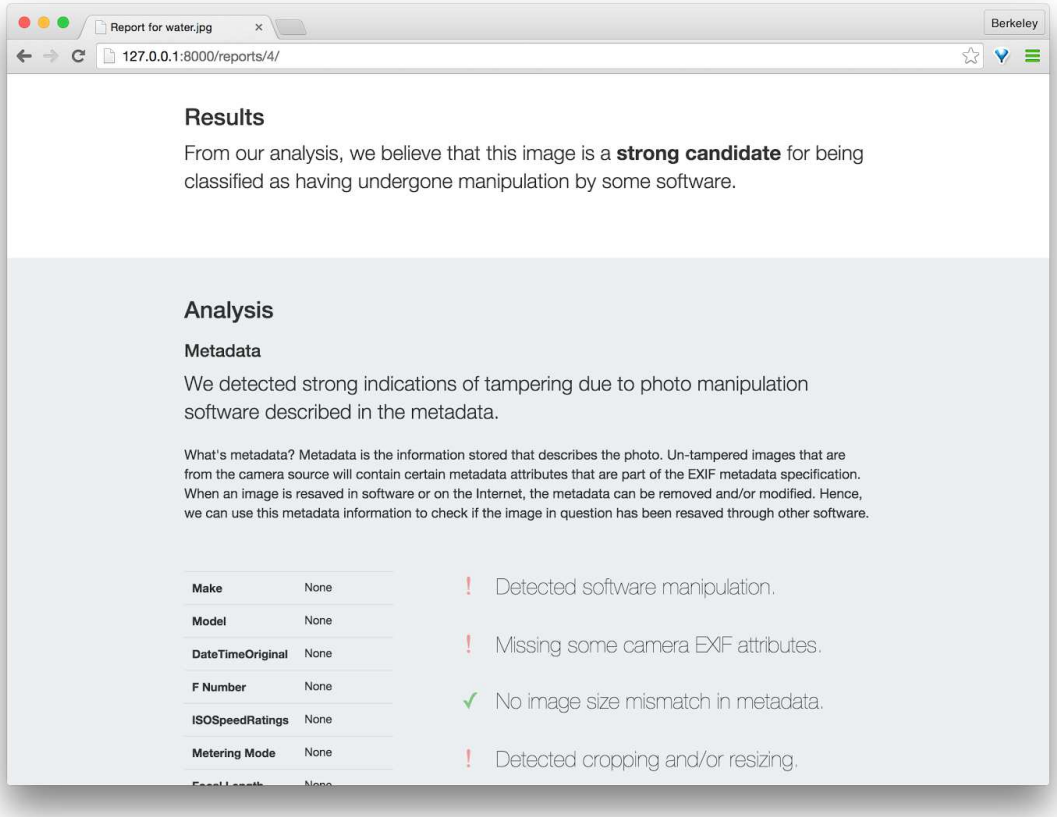
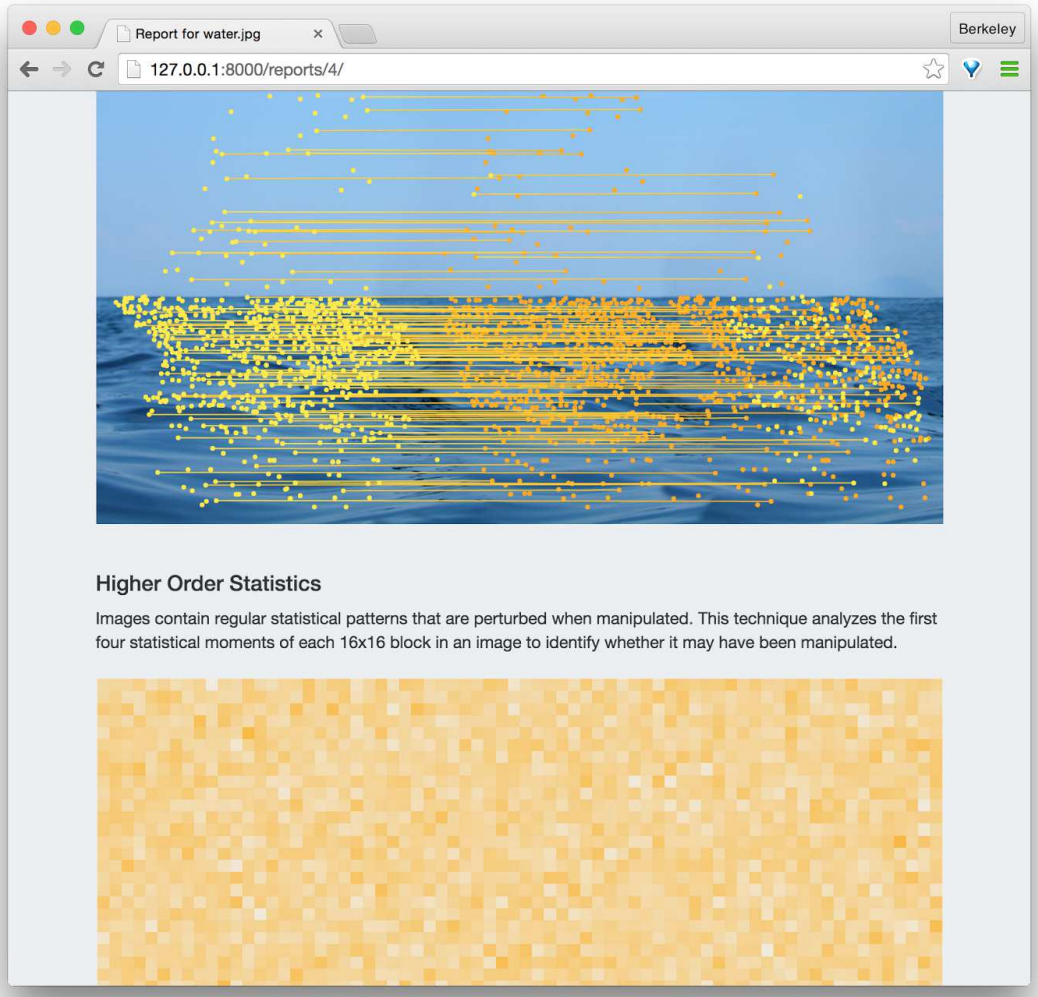


Figure 5: The results page detailing a copy-move detection and higher order statistic analysis.



Individual Technical Contributions

By Anthony Sutardja

Introduction

Photo manipulation tools have become more powerful and accessible to use than ever before. Tools like Adobe Photoshop allow anyone to enhance and alter photos with very little experience. As these tools continue to advance, people will have a harder time distinguishing manipulated images from authentic images.

Our capstone team has developed a set of algorithms to detect manipulated images, and has integrated these varying detection methods into an accessible web service for anybody to use.

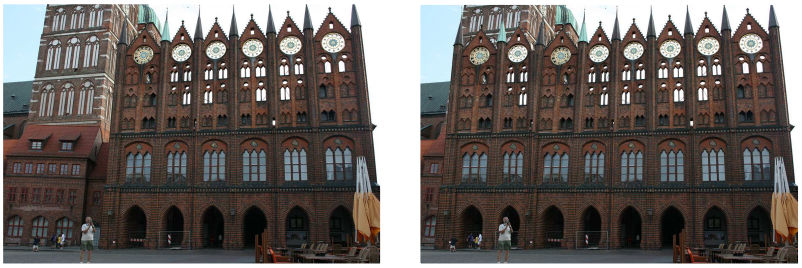
As a whole, there are many layers in which analysis can be performed on an image to determine authenticity. At the most basic level, there is the image's metadata, which can be trivially detected. Images can also be analyzed to detect the compression artifacts left behind from multiple saves of the image. Lastly, there are also the higher-level visual features. Most manipulated images have at least one of the following manipulations: image splicing or copy-move forgery. Image splicing is when a region from another image is patched into a target image. Copy-move forgery is when a region from an image is resampled to another portion of the same image.

Although the team worked on analyzing many of components together, my contributions to the capstone project primarily fall into two significant portions: development a method for copy-move forgery detection and the user interface for the web service. In this paper I will focus only on my contributions towards copy-move forgery detection, since developing and designing user interfaces for the web service is a trivial task.

Copy-Move Forgery Detection

Overview

Copy-move forgery is when portions of the image are resampled to other portions of the image with the intent to change the photo's meaning and context. This resampled region is typically translated, scaled, and rotated. Hence, our technique must be robust to these transformations.



(A) Original Image (B) Copy-moved image

Figure 1 An example of copy-move forgery from the dataset provided by the University of Erlangen-Nürnberg (Christlein et al.). On the left is the original image, and on the right the manipulated image contains a portion of the building copied over.

Some more sophisticated copy-move forgeries use methods like Poisson blending to adjust the lighting, hue, and saturation of the resampled regions to match its surroundings (Pérez et al.). Hence, our copy-move algorithm must also be able to detect manipulations under these different conditions.

The resulting method is a combination of methods that yields reliable copy-move detections in a reasonable amount of time.

Literature Review

There are many different papers that detail existing techniques that differ in their individual ways to attempt detecting Copy-Move forgery.

One of the simplest methods is to perform a sliding window match, where every window is a small subset of the image and is compared to every other part of the image. One technique has made iterations on this concept by sorting lexicographically for faster matching and using quantized Discrete Cosine Transform (DCT) coefficients to perform a more "robust" matching process (Fridrich et al. 8). However, the sliding window method is still significantly flawed in the way in which it functions. Fridrich et al realize that their robust matching technique has a propensity for yielding an alarming number of false positives. Furthermore, although not mentioned by the authors of the paper, an analysis of Fridrich et al's techniques will reveal that these methods are not scale invariant, when an object is rescaled in size, and rotational invariant, when an object is rotated by some angle (Bayram et al. 1055). Both scale invariance and rotational invariance are commonly found amongst Copy-Move manipulations and must be detected.

In order to account for scale and rotational invariance, some techniques use a combination of Harris corner detection and scale-invariant feature transforms (SIFT) to use as a metric for matching across the entire image using a KD-Tree (Shivakumar et al. 2). The drawback, however, is that some of these

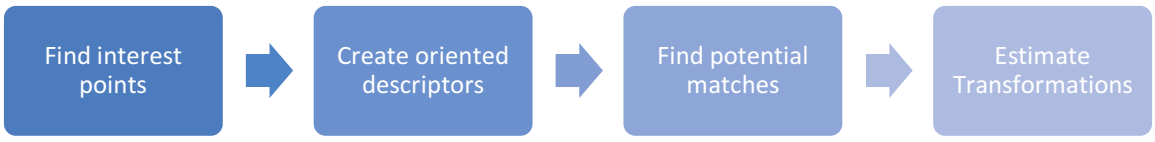
methods have large processing times of up to 2600 seconds (around 45 minutes), which would not be scalable for running analyses on large batches of images (Shivakumar et al. 13).

There also exists other techniques that look at higher-order wavelet statistics to identify forged regions, but these are less applicable to copy-move forgeries in that the copied regions are from the same images and, hence, will be compatible with the rest of the image (Farid and Lyu 6).

The technique that I have developed for our capstone project draws inspiration from methods in panorama stitching, as well as some of the aforementioned techniques. The specific methodologies are described in the "Methodology" section.

Methodology

Our method described in this paper draws inspiration from a technique for stitching together separate images into a panorama. In the panorama-stitching pipeline, the algorithm detects interest points, creates oriented descriptors around the interest points, finds potential matches, and estimates the transformations from the source to destination on a set of inlier points (Brown et al. 1). Rather than stitching separate photos together, our technique identifies key features within the image that have a strong transformation correlation to other portions of the image. In essence, our method attempts to "stitch" the image in question to itself.



Interest Points

In order to even begin matching, "interesting" candidate points are extracted using the Harris corner detector. Harris corners are typically used for interest point identification in computer vision applications, and serve as a fast and rotation invariant interest point (Brown and Lowe 2) (Schmid et al. 158). Here, Harris corners are found by finding points that have high gradient in both the x and y direction of a 3x3 pixel window.

We then select a subset of Harris corners for use with corner-based sorting and Adaptive Non-Maximal Suppression (ANMS).

The reason for filtering the interest points is for speed. Filtering the points for a subset is necessary to drastically decrease running time. Although using every interest point found can lead to very accurate detections, comparing every interest point is computationally expensive and leads to extremely high processing times like seen in Shivakumar's Copy-Move detection time of up to 45 minutes per image. Furthermore, filtering is also helpful in reducing the number of iterations that Random Sample Consensus (RANSAC) needs to perform in order to estimate the various copy-move transformations. For more details about RANSAC, see the section below on "Transformation Estimation."

The two different interest point filtering methods have significantly different characteristics on the detection as a whole.

In corner-based sorting, the top N Harris corners with the highest magnitudes are used, which selects the "best" corners as interest points. For our implementation, we take the top 3000 Harris corners. Taking the top N interest points from corner-based sorting helps capture more localized details. This is especially important with images that have extremely small manipulations, since copy-move transformation estimation requires at least four points in order to come up with an accurate transformation.

In Adaptive Non-maximal Suppression (ANMS), we take the Harris corners that are high in value and relative in spread. This methodology is directly taken from the Multi-Image Matching using Multi-Scale Oriented Patches paper; please see the paper for more details on its implementation (Brown et al. 2). ANMS helps select a broader selection of points that captures more of the full context of the copy-moved manipulation.

Figure 2 Figure 2(A) and 2(B) show the contrast between using ANMS versus corner-based sorting on small regions of manipulation. Here the bird is the manipulated region. Corner-based sorting is able to capture more interest points for matching.

(A) Interest points from ANMS

(B) Interest points from corner-based sorting

Figure 3 Figure 3(A) and 3(B) show the difference between the locality of detection from corner-based sorting and the spread of detection from ANMS.

(A) Detection with top harris corners (B) Detection with ANMS

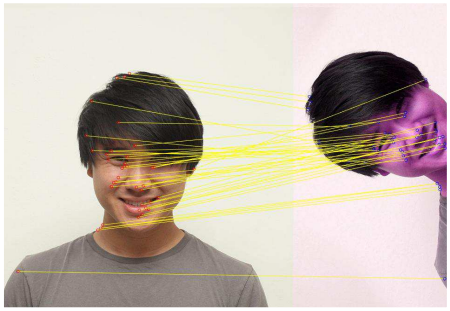
Ideally, our technique should be able to capture small manipulations while also being able to capture as much of the manipulated region as possible. Hence, the interest points used is the union of the interest points yielded by the top 3000 interest points from corner-based sorting and the top 3000 interest points from ANMS.

Descriptors

Once the interest points have been selected, we extract a descriptor for every interest point. These descriptors will be used to match against the descriptors of all other interest points to establish a match.

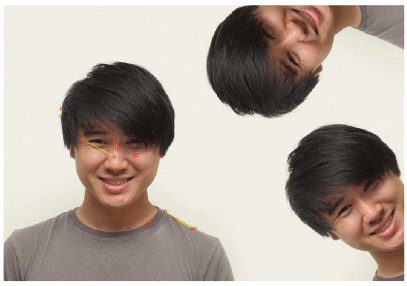
We experimented with two different descriptors: the rotationally invariant box descriptor and scale invariant feature transform (SIFT) descriptor.

Starting with the box descriptor, we first implemented a naive box descriptor that did not take orientation into account. The box descriptor is created by sampling a 40x40 pixel window that is Gaussian blurred (at $\sigma = 1.0$) and then downsampled to an 8 x 8 window, which is subsequently normalized such that the features are zero mean unit variance (Brown et al. 3). The down sampling gives robustness for the descriptor in matching to other descriptors. The normalization also helps to robustly match patches where there are changes in hue, saturation, and lighting. These changes are often applied to help convince the viewer that the copy-move region belongs in the attacked image context.

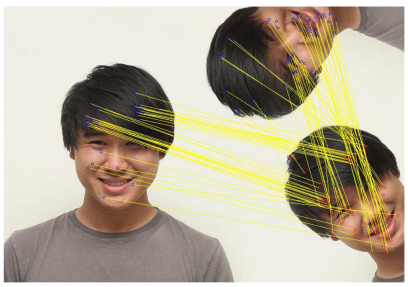


An example of detection of a resampled image with hue and saturation alterations because of normaliation.

Although this descriptor worked well in identifying strictly translational copy-move instances, this descriptor performed poorly for rotated copy-move instances. Hence, the box descriptor was made rotationally invariant by orienting the descriptors such that the dominant gradient direction was aligned with the axis of each Gaussian blurred 40x40 window (with $\sigma = 4.0$). The dominant gradient was found by bucketing each pixel's orientation into ten 36-degree buckets. Each entry into the bucket of angles was weighted by the Gaussian coefficient based on each pixel's distance to the central interest point and the magnitude of each pixel's gradient. The weighted average of all the point orientations that were placed into the largest bucket was used to then calculate the overall weighted orientation of the descriptor. We also included the weighted orientation of buckets that were local maxima (taking the buckets that were at least 0.8 times the magnitude of the largest peak). This technique achieved great results in identifying rotational copy-move as seen below.



(A) With vanilla box descriptors



(B) With rotation-invariant box descriptors

We also analyzed using scale invariant feature transform (SIFT) as a descriptor for interest points, since many different detection techniques also used SIFT for its scale and rotational invariance properties to

detect copy-move forgery (Shivakumar et al.)(Amerini et al.). For our implementation of SIFT, we take a 16x16 window around the interest point and divide it into 4x4 windows. For every 4x4 window, the orientations and magnitudes of the gradients within the window are placed into a histogram with 8 buckets (Sutardja and Tee).

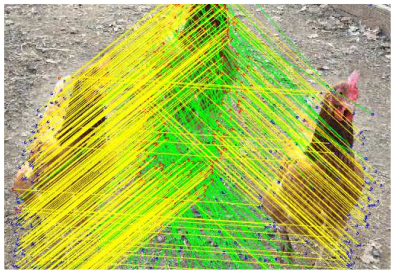
Based on the experimental results that compared the rotationally invariant box descriptor against the SIFT descriptor, we chose to use the rotationally invariant box descriptor instead of the SIFT descriptor for processing manipulated images. Although our results showed that SIFT performed far faster than the rotationally invariant box descriptor, the box descriptor performed for more accurately especially with respect to rotated features. See the discussion in "Results" for further details.

Matching

After the descriptors have been extracted, we use nearest neighbor with outlier rejection to propose matches between corresponding interest points. Specifically, the Euclidean distance metric is used to find all the nearest neighbors for each descriptor. In panorama stitching, the ratio between the first and second nearest neighbors for a given interest point is compared with a threshold value, which determines whether or not the match should be rejected or not due to its non-discernibility from other points (Brown et al. 4). However, this ratio matching called 2NN is non-optimal for images with more than one copy of a source region, since this technique would reject points that have more than one similar descriptor.

Instead, we adopt the technique $g2NN$ as described by Amerini and her group at the University of Florence. $g2NN$ accounts for the high dimensionality of the feature descriptors by noticing that dissimilar descriptors share very high values (Amerini et al. 4). Hence, instead of comparing just the first nearest neighbor with the second nearest neighbor, we match the first g closest nearest neighbors that are under the outlier rejection threshold. This technique enables our algorithm to capture multiple copy-move regions of the same source.

(A) The chicken has been copy-moved to 2 other locations in the image



(B) 395 matches from 2NN displayed in yellow and 251 ***additional*** matches from g2NN displayed in green

In 2NN and $g2NN$, smaller threshold values find stronger matches, and subsequently make running transformation estimation faster. However, in contrast, a larger threshold will increase the number of matched patches, but increase the running time of finding the set of transformation estimations. As suggested by Brown's multi-scale oriented patches paper, we use a threshold value of 0.55 to get both speed and breadth of points (Brown et al. 4).

A minor detail in the matching process is that we must set the distance to infinity for every comparison of an interest point to the same interest point, or else the algorithm would find itself as a match.

To improve the speed of the transformation estimation, we also filter through all the corresponding points to remove the matches that are only a few pixels away from each other. We remove all matches that are only 3 pixels or less apart with the assumption that copy-move regions are spaced at a significant distance.

Transformation Estimation

Once the candidate matches have been proposed as corresponding points, transformation estimation is used to identify the scaling, rotation, and translation to guess the copy-moved region of the image.

Transformation estimations can be achieved with a technique called random sample consensus (abbreviated RANSAC), which acts as a filter for correspondence points that are not correlated with a transformation. RANSAC randomly selects 6 interest points to compute a homography transformation (in our case just an affine transformation) in the hopes that it will produce the best transformation matrix (Fischler and Bolles 3). Once the homography is computed, the sum of squared differences (SSD) is used to filter an inlier set based on an error threshold from all the corresponding points. In every subsequent homography computation, the size of the newly generated inlier set is to the size of the largest

encountered inlier set thus far. If the new set is larger, the algorithm sets the new set to be the largest. This process repeats for a specified number of iterations. Please see the paper by Fischler and Bolles for more details.

The problem with directly using RANSAC for matching correspondence points within an image is the unknown direction of the correspondence points from the transformation source to the transformation destination. Although distinguishing copy-move regions from source and destination is beyond the scope of the paper, the algorithm still needs to be able to "order" corresponding points such that the direction of the corresponding points is consistent with the mapping between one group and another group. If the directions are inconsistent, a substantial number of potentially good corresponding points will not be used in unison with corresponding points that have opposite direction.

Surprisingly, RANSAC estimation works without ordering the corresponding points when the error threshold is set to a large pixel distance (we used a pixel distance of 75). We were able to achieve detections with high recall but noisy precision. These transformations did not seem to be explainable by any affine transformation. However, using RANSAC without directional ordering introduced a noticeable number of false positive matches in each image.

In order to obtain consistent directions among corresponding points, we perform a novel but simple ordering technique that we call "grouping." We select a random point p from the source region that is arbitrarily decided by randomness. For each correspondence (a, b) in the matched corresponding points, we re-order the correspondence (a, b) such that the point closer to the selected central point p is grouped with the "source" region and the corresponding point of s is grouped into the "destination" region (i.e. the points are reordered such that $(min(||a - p||, ||b - p||), max(||a - p||, ||b - p||))$). After grouping is complete, running RANSAC on the arbitrarily grouped "source" points will allow us to estimate a more robust transformation with more possible inlier matches.

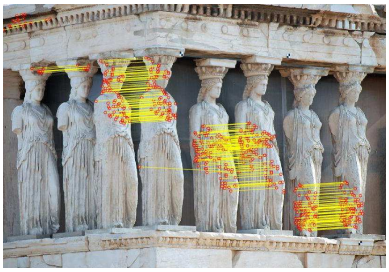
With correspondence points grouping, our algorithm was able to yield consistent detection results after only 15,000 iterations with extremely tight pixel estimation error bounds (set to a maximum distance of 3 pixels). From our experiments, we found that tighter estimation error bounds led to fewer instances of false positive matches.

That being said, vanilla RANSAC did not fit all of our needs for detecting image manipulations. RANSAC is only able to generate a single transformation estimation and does not cope well with multiple

manipulations within the same image. Copy-move forgeries often have many regions resampled from the same image.

We have created a modification of RANSAC that we call multi-RANSAC (mRANSAC). mRANSAC performs RANSAC multiple times to find different sets of corresponding points that each undergo its own affine transformation. In each iteration of mRANSAC, the algorithm first performs a random "grouping" in order to assign a uniform direction for the correspondence points. The algorithm then proceeds with running an instance of RANSAC (with 15,000 iterations), which selects the best inlier set that has at least 7 points. We use a threshold of at least 7 points because we do not want to include the 6 correspondences that were used for the transformation estimation. The inlier set of corresponding points is then removed from the potential corresponding points, and subsequent iterations repeat the process on the reduced set of corresponding points. Once the iterations are finished, we aggregate the inlier points of all iterations and return the matches as our final result.

(A) Copy-move example image from the University of Erlangen-Nürnberg (Christlein et al.)



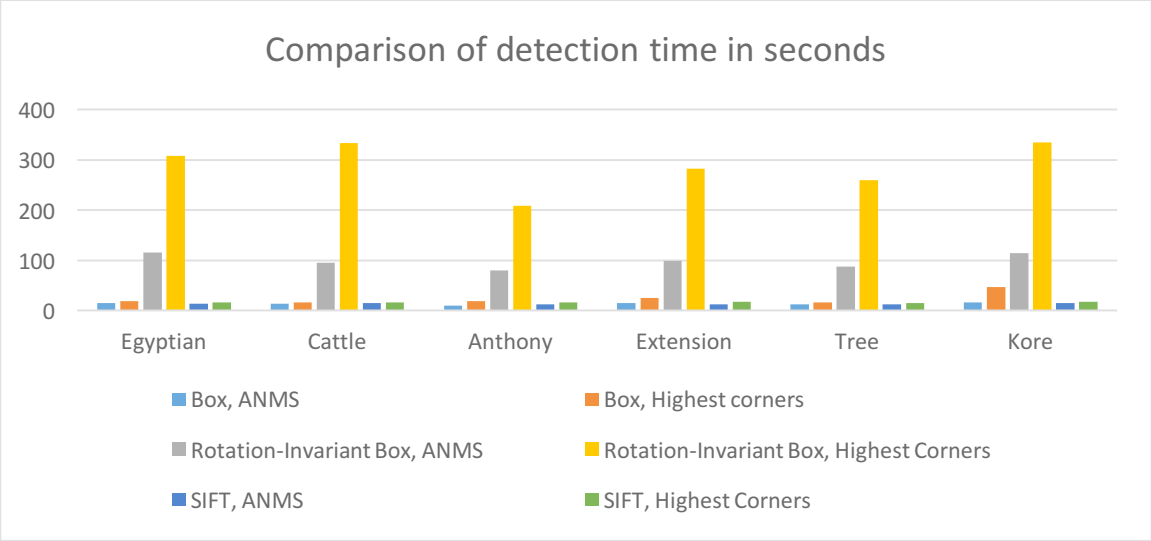
(B) Copy-move detection with vanilla RANSAC and no grouping

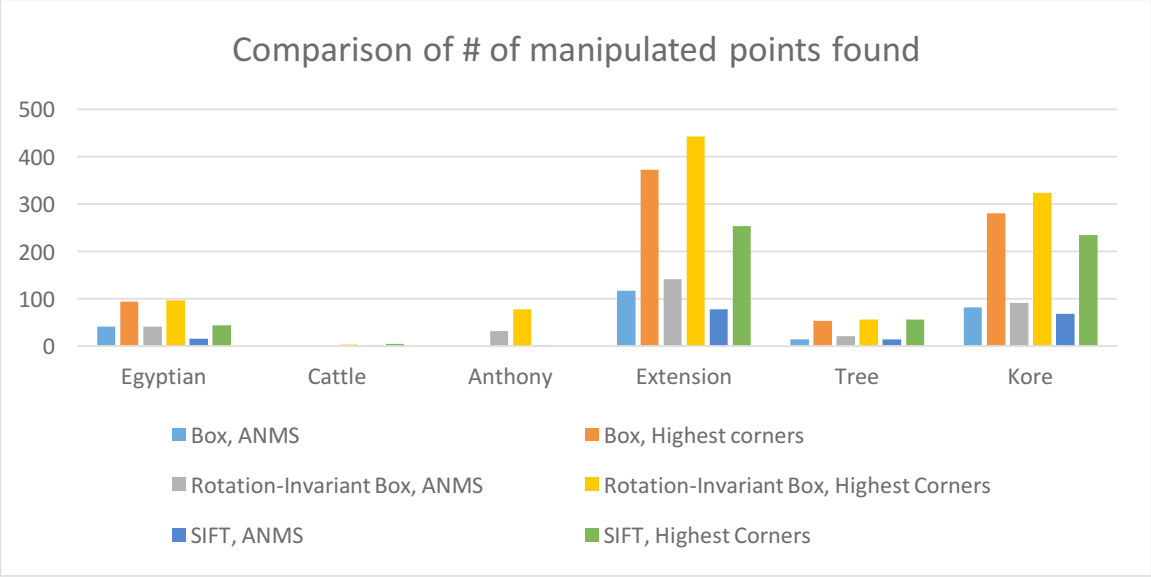
(C) Copy-move detection with mRANSAC and grouping

In our implementation, we run mRANSAC with 5 iterations. The number of iterations directly relates to the number of copy-move regions mRANSAC can detect. Each iterative instance of RANSAC either detects a new copy-move region, or expands an existing detected copy-move region. Hence, we are able to identify more than one copy-move instance within an image.

Results

Our method for copy-move forgery was benchmarked on two metrics: number of points that were correctly identified as copy-move, and the time in seconds for completion. Each combination of techniques was benchmarked on 6 separate images.





Overall, we are very satisfied with our copy-move detection algorithm, as it has correctly identified most instances of copy-move forgery in the University of Erlangen-Nürnberg's dataset (43 out of 48). Our results outperform most other algorithms in either detection time or accuracy. Furthermore, our approach

yields additional information like source to target correspondences. For the majority of the algorithms that we compared our method to, our algorithm was able to identify more copy-move regions with less false positives. For the single algorithm that performed better in accuracy than our method, our method was able to identify a sparser copy-move region in far less time (45 minutes versus 15 seconds with SIFT).

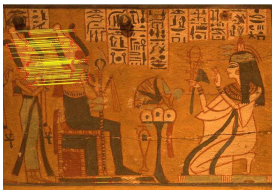
Of the small number of failed copy-move detections, we noticed two problems. First, for images with repeated patterns and lots of symmetry, our copy-move detector had high recall but poor precision (meaning the introduction of a lot of false positives). Many of the symmetrical features in the scene were matched to each other. This could be fixed by adjusting the nearest neighbor outlier rejection threshold to be more restrictive, but this would cause our detector to lose the generality generating enough matching points for RANSAC estimation to function. Second, for the images that we failed to detect copy-move forgery, the copy-move manipulations in these instances were extremely small in comparison to the full image size. Furthermore, these copy-move manipulations were more complex in that the manipulations contained occlusion to help blend the attack into the scene. We can increase the recall of matches within these smaller copy-move regions by using a smaller window size, but this results more ambiguity in distinguishing patches in nearest neighbor outlier rejection. We have made the trade off in choosing thresholds that allow us detect a variety of copy-move attacks at the cost of not being able to pick up small and complex manipulations.

There are two key insights from our work on copy-move detection algorithms. First, there is a trade-off between speed and accuracy. More interest points can lead to more accurate and complete detection, but at the expense of computation time. Using a sparse set of interest points decreases running time significantly, but forces us to choose a strategy in selecting the optimal sparse set of interest points for matching. This leads to our second insight: the two different strategies that we tried have their different strengths. Filtering by Adaptive Non-Maximal Suppression enables our algorithm to capture the overall context of the copy-moved region, while filter by corner-based sorting enables our algorithm to capture smaller copy-moved patches. Hence, using a combination of both of these filtering techniques helps in identifying copy-move forgeries.

(A) Original Image

(B) Copy-moved image

(C) No copy-move detection on original



(D) Copy move detection on manipulated image

From the results, we also see that the rotation-invariant box descriptor helps identify more manipulated points than the SIFT descriptors, but the SIFT descriptors are faster to compute (almost by a factor of 20). That being said, our project requirements don't necessarily require fast computation time, but necessarily requires precision in detecting manipulations. Hence, our deployment version of copy-move forgery detection uses the rotation-invariant box descriptor.

Future Work

One area of future work includes detecting scaled copy-move detections. Currently, our descriptors are not as scale-invariant – meaning copy-move regions of different scales could be missed by our detection technique. However, scaling tends to be less common in copy move examples, as objects can appear out of place with too much scaling and are therefore more easily detected by the human eye.

Towards the end of this project, we found that Irene Amerini and her group in Italy have developed and detailed a copy-move detection methodology that is very similar to our technique (Amerini et al. 1). We have used some of their insights on $g2NN$ that has helped improve detection accuracy for multiple copy-move manipulations. The group also has some interesting techniques in overcoming some of the issues that we face with our algorithm.

First, we recognize that discarding the SIFT descriptor may have been a premature course of action. The SIFT descriptors performed quite well for most images, but poorly for images with multiple manipulations and rotations. Rather than blaming SIFT for poor accuracy performance, the problems in detection accuracy of SIFT descriptors could have very well been due to the nearest neighbor outlier rejection methods that we are using for creating matches. We have only analyzed SIFT's performance with 2NN and have not confirmed the improvements with $g2NN$. Implementing this in our own copy-

move algorithms could yield the accuracy improvements with SIFT that can match the rotation-invariant box descriptor, which would allow our algorithm to perform much faster.

Another area of future work is to further improve mRANSAC and compare it to other techniques in estimating a variable number of manipulations. mRANSAC can be modified such that the iterations continue until there are no sizeable matches, allowing it to detect an arbitrary number of copy-move regions. Amerini and her group have also come across the problem of detecting multiple regions as well and proposed an alternative approach for the transformation estimation. Rather than collecting multiple sets of inlier matches, Amerini's copy-move detection method uses agglomerative hierarchical clustering on the matched points prior to running RANSAC (Amerini et al. 5). Once the clusters are established, the transformation estimation for each cluster is computed with RANSAC. This results in a higher number of detected manipulation points. Incorporating this technique into our copy-move methodology would improve detection robustness.

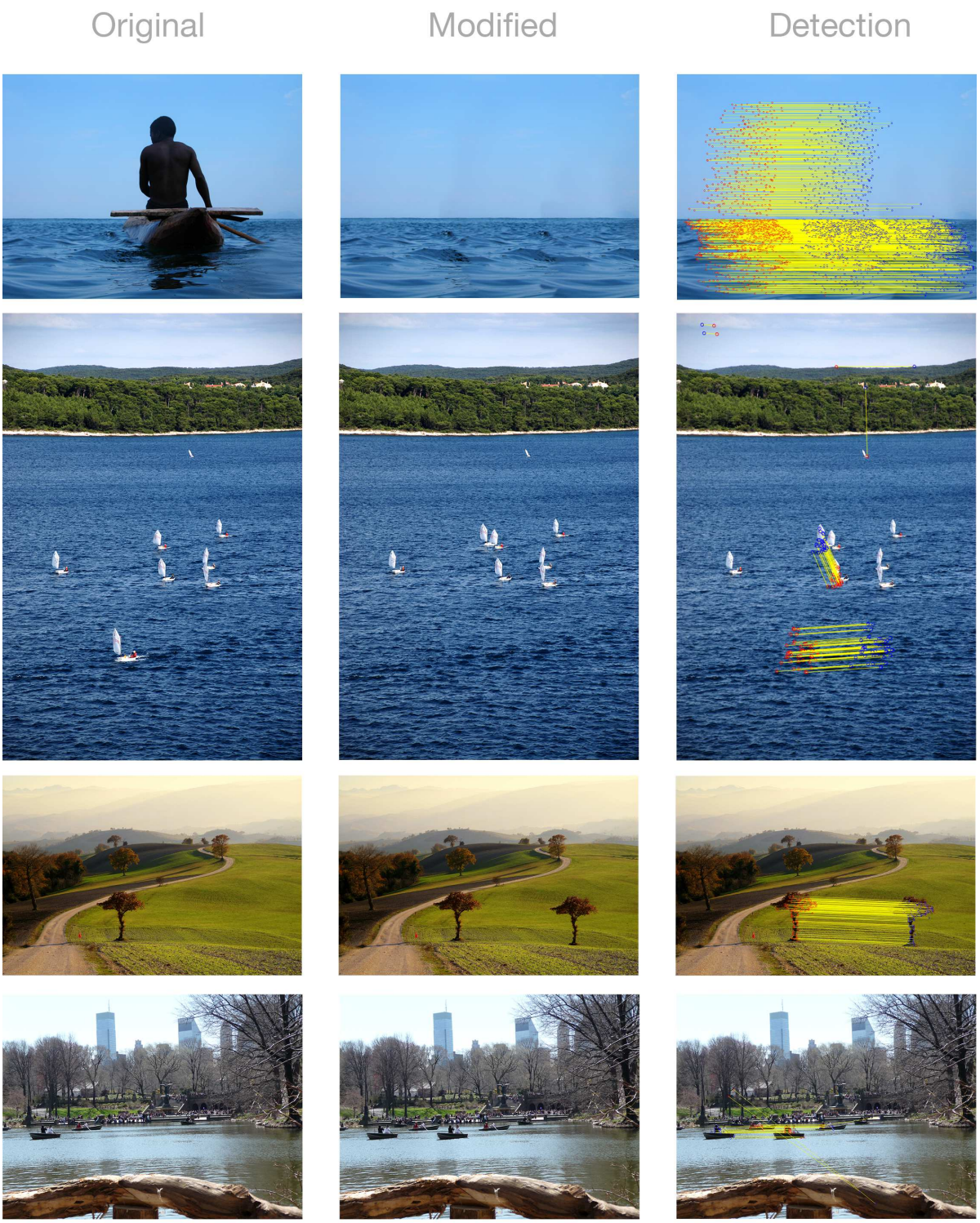
Lastly, the last piece of future work is identifying the region that is the source of all the copy-move resampling. The copy-move detection algorithm described in this paper only finds copied regions without knowledge of which region is the source and which regions are the copies. It is worth investigating the classification of these copied regions into these two classes by using cues like higher order statistics and JPEG error analysis.

Acknowledgements

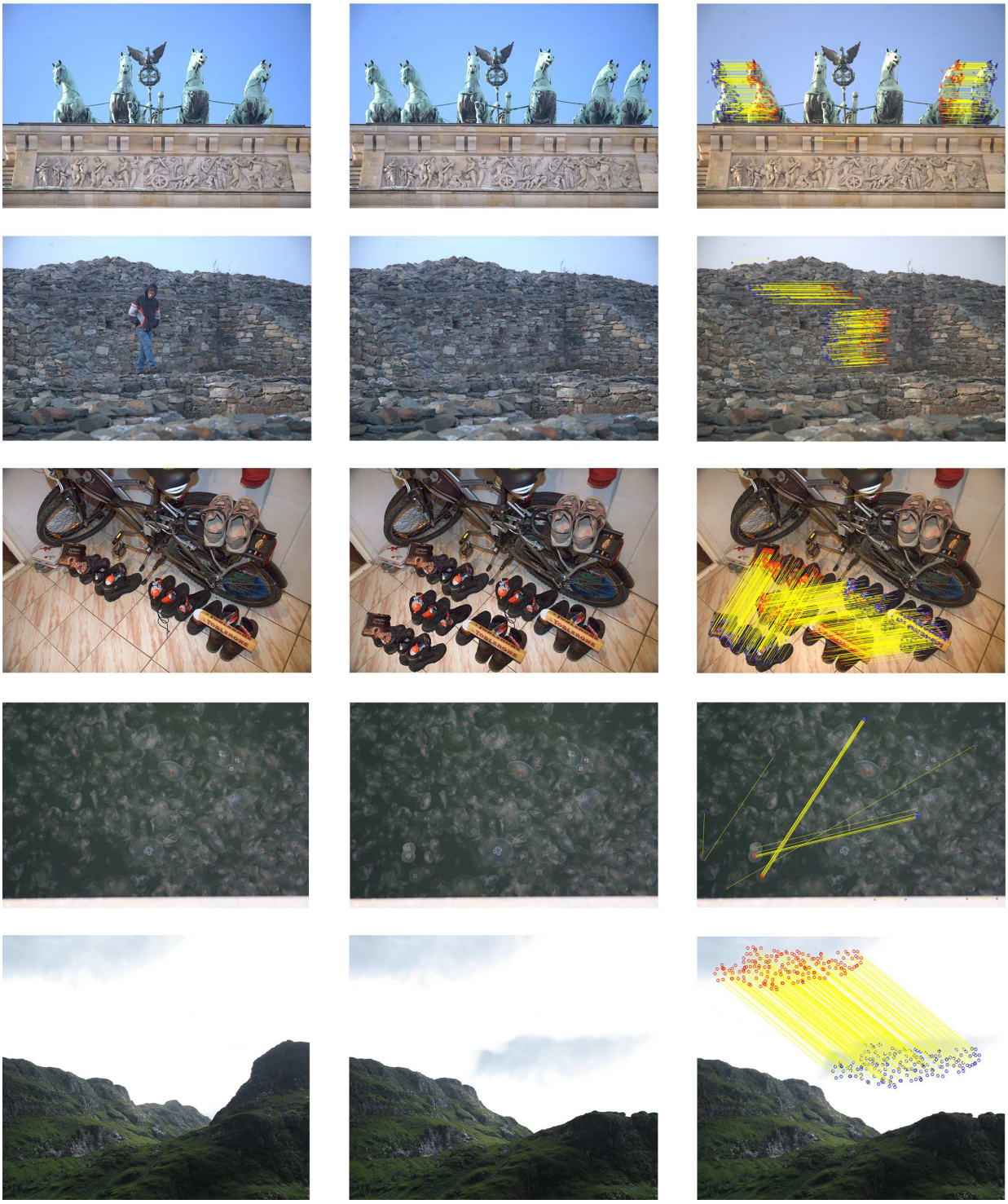
As part of my C294-26 Image Manipulation and Computational Photography class with Professor Aleyosha Efros, I worked with Kevin Tee for a final project on this topic (Sutardja and Tee). I thank Kevin for his development and analysis of SIFT descriptor features. A special thanks also goes to the University of Erlangen-Nürnberg for creation of the high quality dataset of copy-move manipulations (Christlein et al.).

Appendix

Figure 4 More examples of successful copy-move detection on the University of Erlangen-Nürnberg dataset.



Original Modified Detection



Works Cited

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.

S. Bayram, H. Sencar, and N. Memon. An Efficient and Robust Method for Detecting Copy-Move Forgery. ECE Dept., Polytech. Inst. of NYU, Brooklyn, NY. IEEE. 19-24 April 2009

M. Brown, R. Szeliski and S. Winder. Multi-Image Matching using Multi-Scale Oriented Patches. International Conference on Computer Vision and Pattern Recognition (CVPR2005) pp. 510-517.

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.

H. Farid and S. Lyu. Higher-order Wavelet Statistics and their Application to Digital Forensics. IEEE Workshop on Statistical Analysis in Computer Vision. IEEE. 16-22 June 2003.

M. Fischler and R. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Communications of the ACM. Volume 24 Issue 6, June 1981. 1 June 1981.

J. Fridrich, D. Soukal and J. Lukas. Detection of Copy-Move Forgery in Digital Images with. Proc. of DFRWS 2003, Cleveland, OH, USA. 5-8 August 2003.

P. Pérez, M. Gangnet, and A. Blake. Poisson Image Editing. SIGGRAPH '03 ACM SIGGRAPH 2003 Papers. ACM. New York, NY, USA. 27 July 2003.

B. Shivakumar et al. Automated Forensic Method for Copy-Move Forgery Detection based on Harris Interest Points and SIFT Descriptors. IJCA. August 2011.

C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of Interest Point Detectors. International Journal of Computer Vision 37(2), 151-172. Montbonnot, France. 2000.

A. Sutardja and K. Tee. Copy Move Detection. CS294-26 Final Project. University of California, Berkeley. December 2014.

Individual Concluding Reflection

By Anthony Sutardja

Outcomes

Our goal from the beginning was to build different techniques for analyzing image fraud. These techniques were to be used to build a general classifier for flagging suspicious images and to generate a report for a web service.

For most of the individual techniques, we obtained the results we set out for and achieved significant results in analyzing metadata, copymove detection, error level analysis, and double JPEG compression. However, more fine tuning is still needed for other techniques. With higher order statistics, we were unable to extract significant information to help us classify images into the two categories. We were also unable to obtain consistent and generalizable results from the generic classifier that we attempted to flag photos as “authentic” versus “manipulated.”

That being said, we did achieve integrating all of our working algorithms into a web service in an aesthetically pleasing and intuitive user interface.

Product Management

In retrospect, our team needed a designated product manager. From the beginning, our team members agreed to participate in a rotational product management role. This system worked well when someone was clearly designated as the product manager. However, the flexibility in designating a product manager for each milestone caused us to. Our team members are in agreement that if we had more time, we would be able to achieve significantly better results. Having a product manager to really enforce the deadlines that we set for ourselves would have helped us stay on course and given extra time to refine our techniques.

That being said, our product management did identify the juncture where we needed to transition from individual divided tasks to a more collaborative workflow. Initially, we divided our project into three very distinct areas. The distinction of the divisions led to less collaboration with each other since the knowledge domain of each area was unfamiliar to all team members. However, as we became more familiar with the image manipulation terms and techniques, we soon realized that having peer feedback on the techniques we were each implementing gave us clearer insight and new ideas that we may not have

discovered alone. Being able to recognize this juncture and transitioning to more collaboration was crucial to the development of our algorithms.

Future Work

There are many features that need further research and development.

Although my implementation of copy-move detection works with high recall accuracy, there are still many improvements that can be made to the algorithm to improve speed, improve robustness, and identify other interesting characteristics. Please see the end of my "Individual Technical Contributions" paper for my extensive discussion on future work.

The web service that we have developed is currently in a prototype state. The service is very usable, but lacks the features needed for a commercial use-case. In order to have a product that is usable by a customer, more user studies and research is needed with the target customer (the insurance companies) to identify the depth and confidence of analysis that is required. Furthermore, as with any enterprise application, future work on the service will need to implement some security mechanisms to protect customer information and personally identifiable information.

Future Onboarding

Image manipulation detection is a niche technical field. Our team members unanimously agree that the onboarding of domain-specific knowledge was the primary factor that caused the most friction in starting the capstone project. None of us had backgrounds in image processing, image compression, or computer vision. Hence, I recommend future onboarding should require individuals working on the project to have gone through a basic image processing and computer vision class. There are also some key academic papers and reviews that were difficult to discover. Having these resources are essential to understanding the primary concepts and terms that are applied to image manipulation detection. The resources cited in our technical papers serve as essential readings for onboarding.