The Ring

THE JOURNAL OF THE CAMBRIDGE COMPUTER LAB RING

Issue LI — May 2019

Who's Who

Mohsin Ahmad (BA07) has been promoted to Director, Hermes Impact Opportunities Fund at Hermes Investment Management.

Gabrielle Anderson (BA08) has been promoted to Senior Site Reliability Engineer at Google.

Stuart Ashforth (BA16) is working as a software engineer at Jump Trading in London.

Joshua Bambrick (BA16, MEng17) is a machine learning engineer at Bloomberg in London.

Barry Bentley (M PhD18) is an Associate Lecturer at The Open University.

Tom Craig (MA06, PhD09) has been promoted to DevOps Team Lead at FIS.

Tim Granger (PhD04) is Director, Software at Roku and is setting up the company's new development office in Cambridge.

Stuart Grigg (BA17) is a software engineer at Sparx.

Matic Horvat (MPhil13, PhD17) is Head of Data Science at Cytora.

Laura James (MEng00, PhD05) has been named one of the17 new Software Sustainability Institute Fellows for 2019. The Institute is a leading international authority on research software sustainability, working with researchers, funders, software engineers, managers, and other stakeholders across the research spectrum.

Laura also recently joined lowRISC as Head of Delivery. lowRISC is a not–for–profit that aims to demonstrate, promote and support the use of open–source hardware.

Dimitra (Mimie) Liotsiou (BA12) has been awarded her PhD in Computer Science from the University of Southampton, titled 'Measuring the Social Influence of Online Communications at the Individual and Collective Level: A Causal Framework'. She is currently working as a Postdoctoral Researcher at the Oxford Internet Institute, University of Oxford.

Charles McLachlan (MA94) has started Portfolio Executive Growth Academy to help senior executives build a portfolio career of part–time executive roles with growing businesses.

John Messer (BA07) is an investor at Tenzing Private Equity.

Jan Ondras (MEng18) is a Research Assistant at USC Institute for Creative Technologies in The Slovak Republic.

Suraj Patel (BA17) is working for JPMorgan Chase as a software engineer.

Tomas Pfister (BA2010) is Interim Head of Research at Google Cloud AI in California.

Alex Rietmann (MPhil18) is co–founder and CTO at thnk AI.

Gabriela Sklencarova (BA16, MPhil17) is working at Google in Zurich as a software engineer.

Nicko van Someren (MA89, PhD94) has joined Absolute Software as Chief Technology Officer.

Bjarne Stroustrup (PhD79) has been invested as Doctor Honoris Causa at University Carlos III of Madrid (Spain).

Richard Tynan (BA17) works for Facebook in London where he is a software engineer.

Daniele Vettorel (MPhil18) is a PhD student at MIT.

The Ring, Issue LI May 2019
Web: http://www.cst.cam.ac.uk/ring/
E-mail: cam–ring@cst.cam.ac.uk
Tel: +44 1223 763585
Post: William Gates Building, Cambridge CB3 0FD
Published three times a year. Copy deadline for the September 2019 issue is August 10th 2019

The Ring is the journal of the Computer Lab Ring, which is the graduate association of the Department of Computer Science and Technology, University of Cambridge.

Hall of fame news

Bango

For the year ending December 31 2018 Bango narrowed annual pre–tax losses, as annual end –user spend more than doubled and boosted revenues.

Pre–tax losses narrowed to £3.57m compared with £3.9m in 2017 and revenue increased 60% to £6.62m.

End User Spend (EUS) increased 106% to £558.2m from £271.4m in the previous year.

Adjusted losses (LBITDA) improved to £0.87m from a loss of £1.57m a year earlier.

For 2019, Bango detailed two main areas of focus, including growing end user spend by more than 100% year–on–year, and gaining significant new revenue from the $50bn app developers spend on app marketing by offering access to audiences through Bango Marketplace.

'App developers, app stores, merchants and payment providers are crossing the threshold into the Bango ecosystem to collaborate, grow and thrive. More mobile commerce throughout the Bango Platform will deliver value from both the established payment platform and also from the unique ability to securely and safely monetize anonymized payment data. This is the opportunity that Bango will focus on during 2019', said Ray Anderson, CEO of Bango.

Bromium

Bromium has announced an expansion of its relationship with HP Inc with the integration of its flagship product, Bromium Secure Platform, into the new HP Device as a Service (DaaS) Proactive Security service.

Bromium's application isolation technology powers HP Sure Click Advanced, which protects endpoints from malware introduced through email attachments, infected links, web browsers or downloadable files.

The HP Sure Click Advanced technology enables the HP DaaS Proactive Security Service to deliver an isolation security service for files and browsing on Windows 10 PCs regardless of whether the endpoints come from HP or another manufacturer.

Calipsa

Innogy Innovation Hub has completed seed investment in AI powered CCTV monitoring company Calipsa.

The investment will enable Calipsa to further refine its product and support its growth strategy in the UK and Europe.

Cambridge Spark

Cambridge Spark, leading provider of data science apprenticeships, training and recruitment, has been awarded £1.4m in funding from Innovate UK.

The funding supports two projects, including a research collaboration with the University of Cambridge, to enhance the capabilities of their proprietary patent–pending AI–powered EdTech platform, the Knowledge Assessment Teaching Engine (K.A.T.E®).

'K.A.T.E.® was conceived with the idea of delivering a scalable software platform, bridging the gap between higher education and industry standards by making personalisation the core of the user's experience', said Raoul–Gabriel Urma, CEO of Cambridge Spark.

Dr Urma said 'It provides instant feedback and grading on code written by data scientists and developers within an industry–simulated environment. This new funding from Innovate UK enables Cambridge Spark to develop support for multiple programming languages, performance and profiling metrics, personalised recommendations, and assessment utilising cutting–edge AI approaches.'

DisplayLink

DisplayLink®, the leading provider of USB graphics and wireless virtual reality technology has announced a commitment to a new 60000 square foot engineering design centre on the Cambridge Science Park to accommodate up to 500 employees.

DroneDeploy

DroneDeploy customers conducted more than 1m automated drone flights in 2018.

DroneDeploy's software platform has helped customers map more than 40m acres across 180 countries.

2018 was the company's most successful year thanks to the surge in drone adoption by the construction sector, the fastest growing sector for commercial drone adoption. DroneDeploy's customers include Jacobs, Skanska, Sundt Construction, Brasfield and Gorrie, Layton Construction and McCarthy Building Companies.

The Federal Aviation Administration (FAA) forecasts another year of extraordinary growth as regulations ease the burden for industry to collect aerial data at scale.

Green Custard

Green Custard has reached the SME Cambridgeshire Business Awards final and is up for Business of the Year 2019.

The final will take place at the Imperial War Museum on July 4th 2019.

Improbable

Improbable has started the development of its own online multiplayer games, powered by SpatialOS, with game development in two studio locations.

Jagex

Old School Runescape has won the EE Mobile Games of the Year at the 2019 BAFTA Games Awards.

Jagex launched the title on iOS and Android in October 2018, and won the award in a public vote which included competition from five other mobile games, including Fortnite, Pokemon Go, and Clash Royale. Since its launch on mobile, the game has been downloaded and installed on 6.4m devices.

Masabi

Masabi has announced that Prestige, the Italian tour bus operator, has launched a new mobile ticketing application 'Enjoy Bus Rome' for the Hop–on–Hop–off sightseeing bus service in Rome.

The app includes the full range of fares and removes the need to visit a travel shop to purchase a ticket. The application is currently available in English but soon will also be launched in Italian, French, Spanish and Japanese. It can be downloaded from the App Store and Google Play.

The new service uses Masabi's Justride ticketing platform, along with the Justride Inspect App for ticket checking onboard and the Justride Hub, a secure cloud–based back office providing real–time data, reporting and analytics, as well as customer service tools.

Spektrix

Spektrix, the UK's leading provider of cloud–based ticketing, marketing and fund-raising software for the arts, has raised £5m from Foresight VCT plc and Foresight 4 VCT plc, to accelerate product development and support Spektrix's international expansion, particularly in North America. As part of the investment, Foresight has taken a minority equity stake in the company.

Spektrix has been recognised as one of the UK's fastest growing technology companies by the Sunday Times TechTrack 100 and FT Future 100, and in 2018 helped almost 400 organisations to sell more than £500m of tickets.

Job listing

April 2019

KTP project
- *Associate*

lowRISC
- *Hardware Design Engineer*
- *Hardware Test and Verification Engineer*
- *LLVM Compiler Toolchain Engineer*
- *Secure Hardware Design Engineer*
- *Senior Hardware Design Engineer*
- *Software/Hardware Design Engineer*

GreenCustard
- *Senior Mobile Engineer*
- *Software Developer*

The Engineering Company
- *Back–end Software Engineer*
- *Fullstack Engineer*

PolyAI
- *Frontend Engineer*
- *Fullstack Engineer*

re:infer
- *Backend Engineer*
- *Machine Learning Research Engineer*

Cydar
- *Computer Vision Engineer*
- *Senior Computer Vision Engineer*

Flourish
- *Developer intern*

Sartorius–Stedim Biotech
- *Software Engineers*

If you have a job advert that you would like included in the weekly listing, please send the details (as a word doc) to cam–ring@cst.cam.ac.uk

Hall of Fame Award Winners 2019

Company of the Year: PolyAI

PolyAI aims to revolutionise call centres with its state–of–the–art conversational AI. The founders of this London start–up met at the Computer Lab while working on their PhDs. Unlike other customer service chatbots, PolyAI's AI technology is able to follow a conversation and interpret meaning according to context, producing more authentic and effective interactions.

Product of the Year: Pur3 Ltd for Pixl.js

Pixl.js, a smart, wireless display, which uses Bluetooth Low Energy, is Gordon Williams' latest creation. With the Espruino JavaScript interpreter it only needs tiny amounts of power to run.

Gordon Williams with the Product of the Year award and a Pixl.js

Espruino was created by Gordon in 2012 and, after a successful Kickstarter campaign, was made open–source in 2013.

Better Future Award: Gemma Gordon

Gemma received the Better Fuure Award for her work on bridging virtual reality with climate change education

Better Future Award winner Gemma Gordon

Gemma Gordon works on the LABScI Imagine project, which aims to create virtual learning environments based on US national parks. This gives children who cannot access nature, such as hospitalised paediatric patients and students of limited means, the opportunity to interact with it through cutting–edge technology delivered via readily–available Google Cardboard VR headsets.

Lessons cover topics with a STEM focus, including: plate tectonics; erosion; desert adaptations; local flora and fauna; human geography and anthropological history. The Imagine project seeks to provide environmental and climate change education, and enable students, who cannot go on field trips, to understand their connection to the natural world.

Publication of the Year: Noa Zilberman, Gabi Bracha and Golan Schzukin for 'Stardust: Divide and conquer in the data center network

The paper presents Stardust, a scalable fabric architecture for data centre networks. It separates the simple network–fabric from the complex network–edge. Stardust applies system–switch architecture on a data centre scale, while attending to the scalability limitations of network devices: resources, I/O and performance. The resulting network fabric devices are a radical change from commodity Ethernet switches, eliminating significant overheads in DCNs. The approach is practical, power–efficient, cost–effective, scalable, and, critically, deployable.

Publication of the Year Award winner Noa Zilberman

To see previous Hall of Fame winners go to https://www.cst.cam.ac.uk/ring/awards

Part IB Group Projects

Juraj Mičko reports on Team Bravo's project Bone Doctor, winner of this year's Technical Prize.

Background

Diagnosing bones using x–ray images can be problematic. Examining x–ray images is a time consuming process and in some cases, medical practitioners can overlook problems and diseases in their first examination.

The task of this project was to implement a platform to be used after the x–ray image is taken and before it is examined by hospital staff. We were asked to provide tools for help and guidance when making decisions, in order to enable radiologists to make decisions faster and more accurately. Therefore, the purpose of the project was set to enhance the decision–making process, not replace it.

The MURA dataset

MURA (musculoskeletal radiographs) is one of the largest public radiographic image datasets of bone x–rays. Contained are 40,000 images of x–rays bundled in 14,000 different studies. The data labels comprise the body part (hand, arm, shoulder, etc.) and the abnormality; whether the bone is healthy (normal) or not healthy (abnormal).

Figure 1: The diversity of the MURA dataset (The MURA dataset is released under Stanford University School of Medicine MURA Dataset Research Use Agreement.)

Features

Refine the x–ray

Usually an x–ray image is far from perfect. Sometimes it contains a hardly visible bone of various greyscale levels or the frame of the physical x–ray image that was scanned to produce the digital image. We implemented a pre–processing phase to normalise the x–rays in the dataset and the x–rays input by the user, with the intention of producing an image of a uniform format which could be more easily analysed by the medical practitioner and more suited to subsequent algorithmic processing.

1. Ensure that the background of the x–ray is dark, and the bones are relatively lighter, inverting parts of the picture if necessary
2. Identify and remove the borders of the scanned x–ray
3. Adjust/rescale greyscale levels to increase the contrast between the bones and the background
4. Crop the image in order to focus on the bone itself

By studying the image and constructing a histogram of the frequency of each colour, the algorithm detected the greyscale levels of the background colour.

Using Sobel's edge detection, the points of the maximum gradient were identified in the image to construct a convex hull of the x–ray using Graham's scan. By using the histogram of colours within the hull, black and white thresholds were established. This range of greyscale levels was re–mapped onto the full range from black to white. The hull was then used to remove borders by drawing a black band along the convex hull, and the image was finally cropped by taking the minimum bounding-box of non–black pixels.

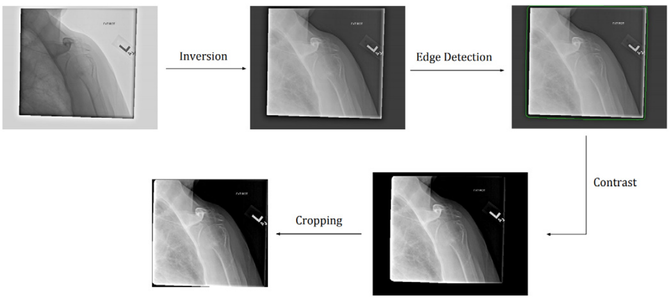

Figure 2: The pipeline of pre–processing of images

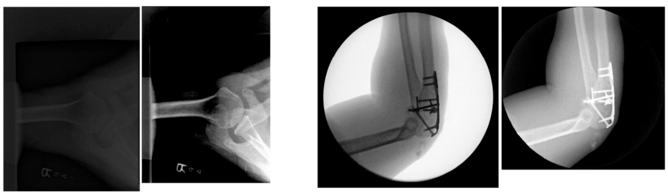

Figure 3 Pre–processing of the left picture of each pair results in the picture on the right, respectively. Note that in the first case, the program turned an almost black contents of the image into a visible bone.

Find a fracture

Classification of normal/abnormal

A binary classification model was built using Keras to distinguish normal and abnormal x–-ray images. The model used in the MURA paper (2017 Rajpurkar et al.) was implemented, the main component of which was a 169–layer DenseNet (2016 Huang et al.). The final layer of the DenseNet was replaced by a single–output fully connected layer with a sigmoid activation function.

Experiments were done with training parameters including learning rate, number of epochs, batch size etc. The parameters that yielded the best results were as follows:

- The model was initialised with pre–trained weights from the ImageNet dataset.
- The model was trained using Adam optimizer with a learning rate of 0.0001.
- The loss function used was weighted binary cross entropy.
- The training was performed via Google Colab, a free cloud computing service with GPU support. Due to the associated RAM constraint of around 12 GB, each body part had to be trained on individually, one after the other.
- The GPU memory constraint allowed for a batch size of 11. Three epochs were run for each body part's subset of the MURA dataset before moving on to the next body part. Six of these 'body part –wise' iterations through the MURA dataset were performed essentially yielding 18 epochs of training in total on the entire dataset.
- The model was then evaluated on the validation dataset and achieved an accuracy of 70.2%.

Figure 4: Example of an x–ray classified as abnormal because it contains a fracture

Abnormality highlighting

The goal of this stage was to help radiologists visually identify where a fracture or other bone abnormality could potentially be located. For this, we used an x–ray image of a similar view of the same body part from a healthy patient i.e. without abnormalities in that bone. How to find such an image from the MURA dataset is explained in later sections.

Highlighting abnormalities was done by hashing each pixel and its neighbourhood in the input image and the similar image. For each pixel in the input image, we looked for similar neighbourhoods in the image of the healthy bone and thus determined what to highlight. Abnormalities in the input image would likely contain pixels whose neighbourhoods were not present in the image of the healthy bone. In some cases, this worked well automatically but in others, the user was required to tune two parameters determining the similarity threshold.

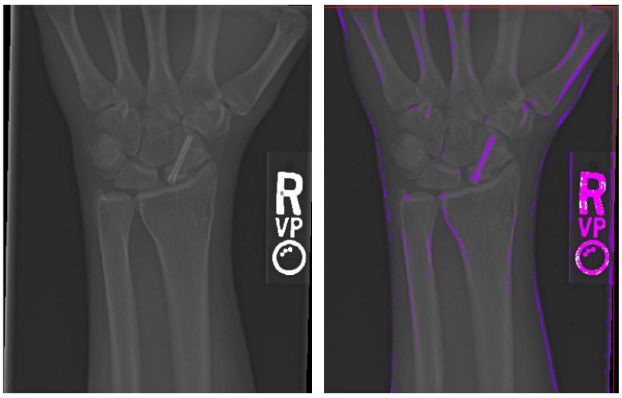

Figure 5: Bone of a hand with abnormalities highlighted

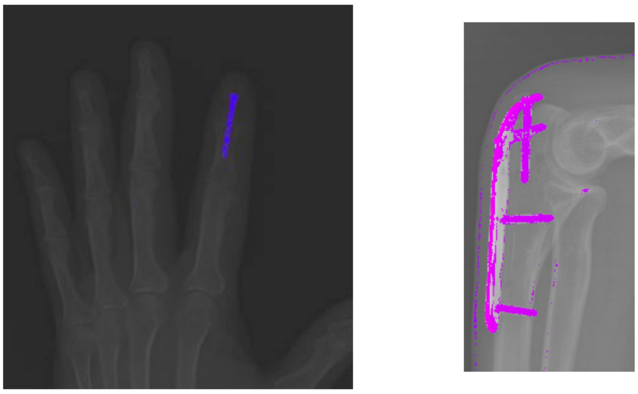

Figure 6: More examples of bones with abnormalities highlighted

Comparison with similar cases

On top of the previous two features, we aimed to help the radiologist diagnose an x–ray in an even more advanced way. When investigating a section of the x–ray with potential abnormalities, it is not always clear whether the bone is healthy or whether an abnormality is present.

To help distinguish between the two outcomes, our goal was to find similar x–rays of healthy bones of the same body part, to enable the radiologist to compare the x–ray being investigated with other healthy cases that were visually similar. Differences between the bone being diagnosed and other healthy bones would help determine whether the case under review was indeed healthy.

Furthermore, we managed to find a mapping between pixels of the input image to pixels of those with similar matches, so that when investigating a specific part of a bone, we could come up with visually similar cropped areas of the same bone part, produced from the set of matching healthy bones.

Step 1. Clustering of images

In order to facilitate the downstream processing of the images for greater performance, the images of the MURA dataset were filtered by finding the views in which they were taken i.e. filtered based on how the bone was situated or oriented when the x–ray was taken. Since the MURA dataset has no labels on the visual view, an unsupervised machine learning approach was taken. Unsupervised learning in this context is the most scalable way to perform this task, instead of handcrafting feature detectors or manually labelling the images and then performing supervised learning. Specifically, we wanted to find clusters of images such that in each cluster, the bones were situated in the x–ray image similarly. Then, given a new image, we could ascertain which cluster it belonged to, and so apply the following algorithms on these filtered set of images belonging to that cluster.

An InceptionV3 network was used to act as an image encoder that took in an image and produced high–level features of the image. This extracted the most significant features and represented them in a vector form suitable for clustering. Congruent with standard research practice, first we used model weights pre–trained on the ImageNet dataset for object recognition, which gave us a model capable of understanding high–level features such as edges, colours and textures. Subsequently, for each input image, we took the mixed_10 layer output of the network, and ignored the last fully–connected layer output. This produced a 8 x 8 x 2048 shaped output tensor, which we subsequently reshaped into a vector of shape 131072 x 1 representing the high–level features of this one image.

Through performing this inference step for all images, we effectively obtained an (N, M) matrix, where N represents the number of all images used and M = 131072 represents the dimensions of each feature vector. Standard K–means clustering is then applied to cluster this feature matrix, producing 2 distinct clusters.

Finally, in order to know what cluster an image belonged to, we performed inference on the image to extract its high–level features and compared the L2 distance to the centroids of the clusters obtained. The closest L2 distance to one cluster label determined which set of images the new image belonged to.

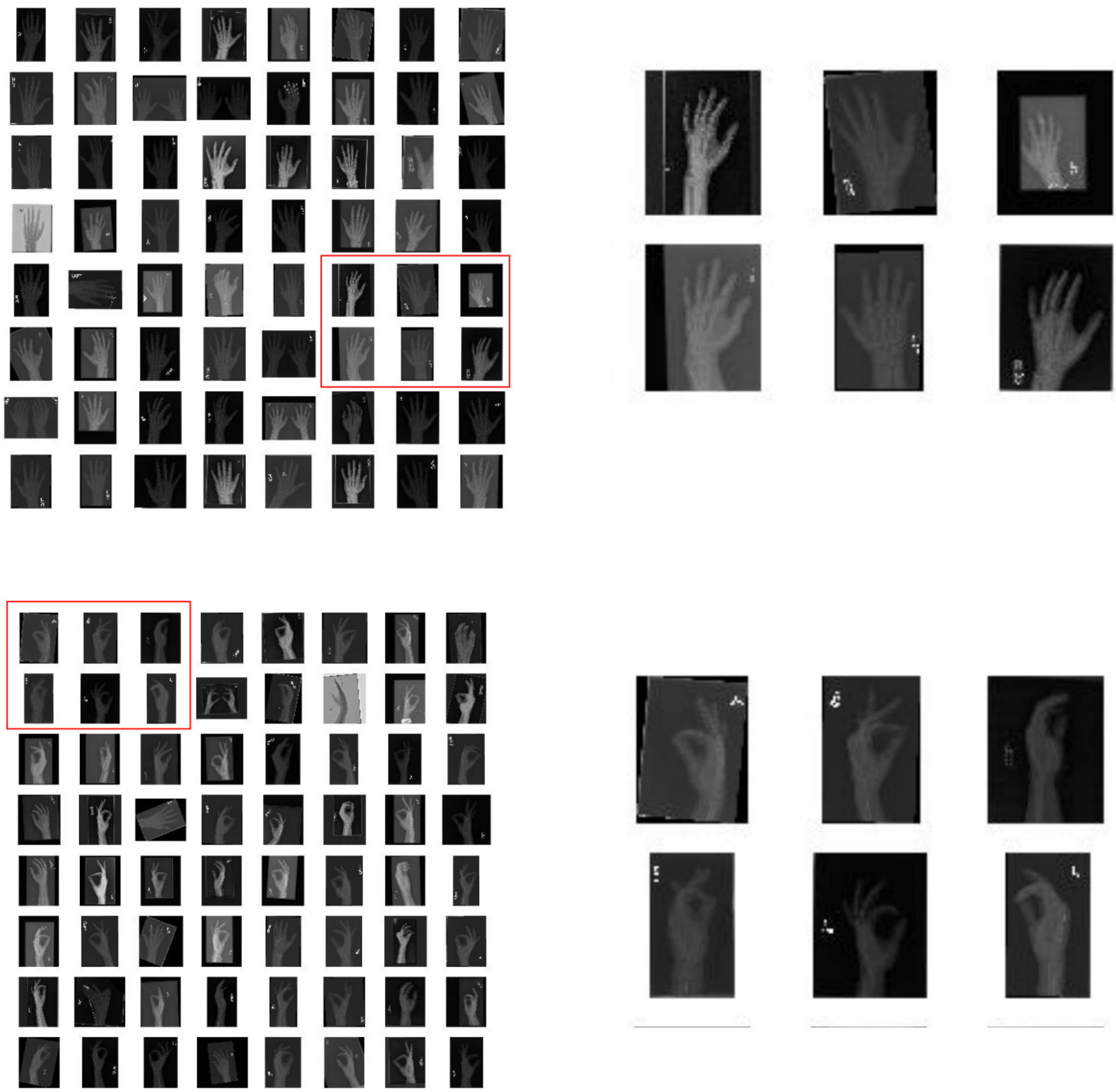

Figure7: Preview of two clusters of hand x–rays

S

Step 2. Image hashing

Given an input image and its cluster produced in the previous step, we wanted to quickly find visually similar x–rays from the cluster. Due to the high number of available images in the dataset, it was infeasible to perform a computation–intensive evaluation on all images in the cluster. The approach we took was to algorithmically pick a smaller number of images that were likely to be similar to the input image, and to use those to further choose the best matches by a more precise evaluation.

Using a perceptual hash, the program compressed the images into a very low–resolution format that could still be searched for image similarity. After hashing the input image, we compared the hash with pre–computed hashes of images in the same cluster in the dataset. This way we could pick a few tens of the most similar images from the dataset, i.e. images with the lowest Hamming distance, in a matter of milliseconds.

Although the hashing algorithm was too primitive for the purposes of finding x–rays of bones of clinically similar cases, it proved to be enough to narrow down the set of images to be used from thousands to tens.

Step 3. Searching for matching images

A precise image matching algorithm was created to further refine the matches obtained, based on visual similarity. As each image had a corresponding encoded image vector, we found the visual similarity amongst images by computing the cosine similarity between any two vectors. Through computing cosine similarity, we could compare how similar the image vectors were in the high dimensional space they were encoded in, with 1.0 being the best score and –1.0 being the worst.

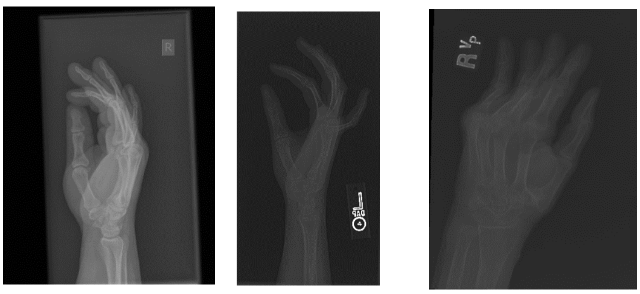

Figure 8 The first two images above show a close match based on having the best cosine similarity of selected images provided from the previous step. The last image shows a bad match with lower cosine similarity despite being in the same cluster. Thus, using cosine similarity is a good indicator of the visual similarity of images using their feature vectors.

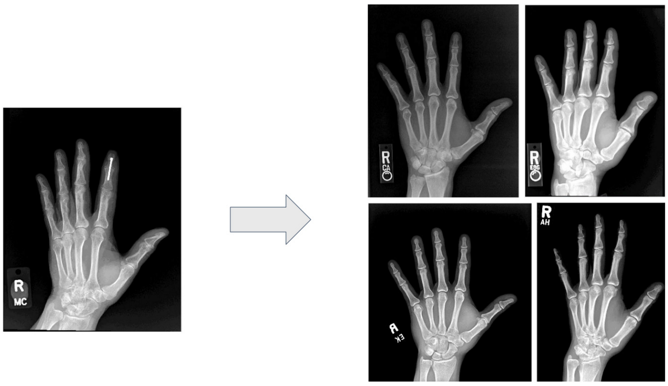

Figure 9 Preview of an input image and a set of most similar images from the dataset, found by the image matching algorithm described in the three steps above

Step 4. Mapping of pixels

The final challenge was to find a mapping from pixels in an input image to pixels in a sample similar image. When diagnosing a specific part of a bone in the input image, this helps the radiologist find and focus on the same area in images of healthy bones of other patients in the dataset. Ultimately allows the practitioner to automatically preview and compare multiple bones, providing more informed and faster diagnosis.

'Base image' refers to the x ray image that was being diagnosed and 'sample image' refers to an image containing a visually similar bone that we wanted to map the base image pixels to.

Mathematically speaking, finding a mapping between pixels of two images is equivalent to finding a transformation of the sample image onto the base image. The reverse of this transformation then defines the mapping. Thus, we tried to visually deform the sample image to match and overlay the base image as much as possible.

First, we defined a set of transformations, such as affine transform (includes translation, rotation, scaling, skewing), and warp transform (using a 6x6 grid with outer points fixed in their original position). In total, these transformations have 38 degrees of freedom, i.e. 38 parameters. Given values of all parameters for this set of transformations, the transformations were applied sequentially to the sample image to produce a transformed sample.

Secondly, this transformed sample image was then compared with the base image, pixel by pixel, each pixel's colour compared with the colour of the corresponding pixel in the other image. The comparison assigned a score to the result of the overlay of the two images — the better match, the lower the value. The scoring function also considered the plain values of the parameters to penalise deformed images more.

Combining the two above, we composed a function that took a set of parameters and returned a score. The program then utilized the Powell's BOBYQA minimisation algorithm to minimise this composite overlay function and returned the parameters of the best transformation found.

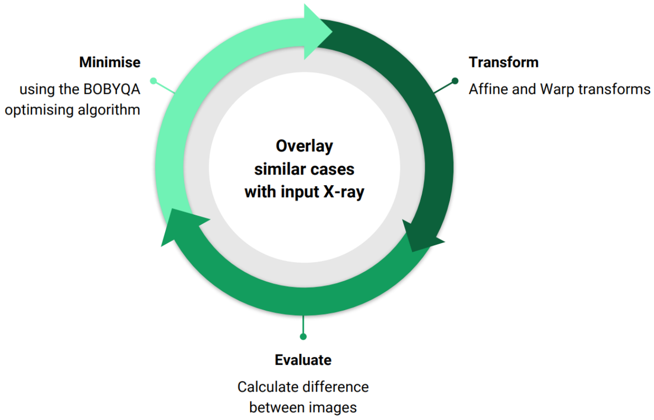

Figure 10 Diagram of the overlay algorithm that uses image transformation and a scoring function to find an optimal mapping between pixels of the base and the sample image

We considered using a Gradient descent optimisation, however, it would very difficult and computationally intensive to compute the gradient of the function to minimise, given a raster image containing hundreds of thousands of pixels. The BOBYQA algorithm only uses function values to minimise the function, with the drawback that it needs more computing cycles to find the optimum. Each overlay algorithm has multiple hardcoded parameters, such as thresholds, scales and multiplicative constants. These parameters were fine–tuned to bring the best results for the majority of images.

In our algorithms, there is a trade-off between performance and precision. We allowed the practitioner to input a 'precision' parameter. Decreasing the precision parameter caused the input image to be downscaled for faster computation and increased the 'Stopping trust region radius' parameter of the BOBYQA algorithm, which specifies the desired absolute precision. The drawback is lower accuracy and a worse match between the original base and sample images.

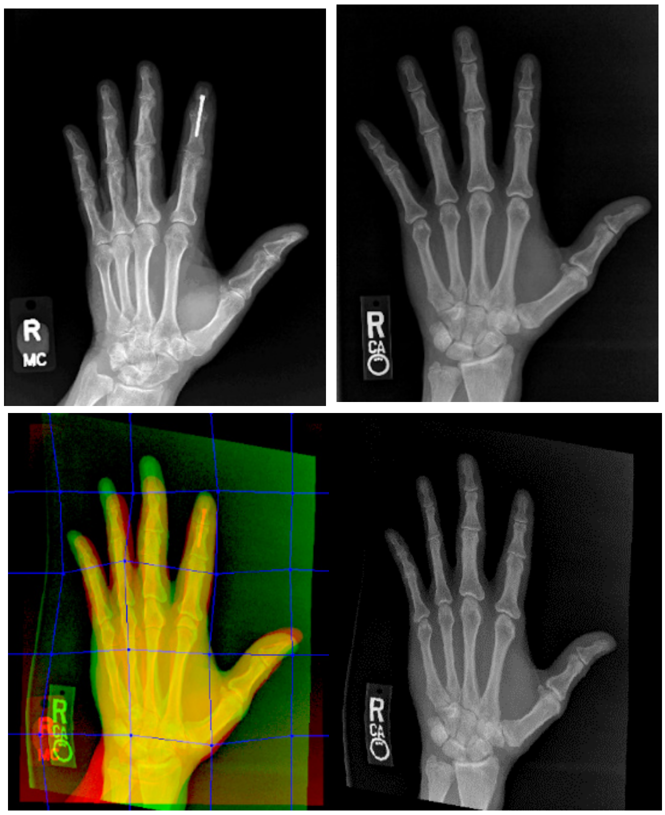

Figure 11 In the following order: base image (top left); sample image (top right); overlay with the base in red and the sample in green; sample image transformed (bottom right) using the found optimal transformation. The pixels in the last (transformed) image positionally correspond to pixels of the same part of the hand in the base image

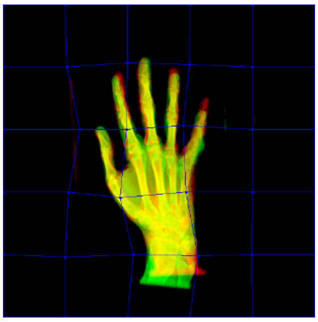

Figure 12 Another example of a visual overlay of a base and a sample image, used to demonstrate the capability of the algorithm to find a perfect mapping of pixels between two x-rays of the same body part. The blue marks are the grid points of the warp transform used.

Connecting the components

The program was written in Java and Kotlin to allow multiplatform use, emphasize robustness, and introduce modularity and extensibility.

Scripts for training neural networks were written in Python. The Gradle build automation tool was used for managing project resources and libraries.

In order to integrate machine learning models built and trained in Python into the Java pipeline, the following was done:

- The TensorFlow graph associated with the model was frozen in Python yielding a Protobuf file.
- The TensorFlow Java API was used in the pipeline to reconstruct the model's graph at runtime from the Protobuf file, feed an input image through it and fetch the classification result.

All the components were connected in a complex computational graph with the property that whenever a user input changes, nodes that depended on it were recursively recomputed. Over 100 nodes were generated when the program was run and so in order to harness computer resources more efficiently, we implemented a parallel computation of the graph where possible.

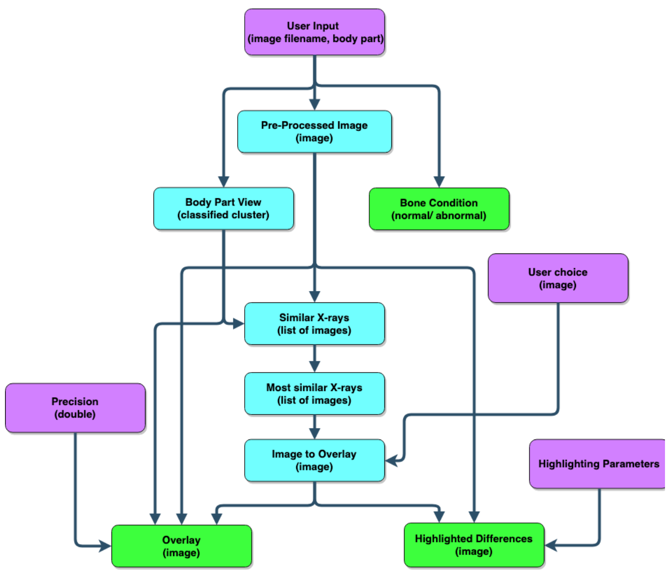

Fig 13 A simplified visualisation of the computational graph

User interface

Although our group focused primarily on the algorithmic parts of the project, we built a user interface to demonstrate the proof of concept.

Starting using requirements analysis, we arrived at a minimum set of features that needed to be implemented. These included:

- X–ray selection screen: to select an image from the local file-system to start the analysis and image matching process
- X–ray analysis result screen: to display the analysis result (normal/ abnormal), a confidence level, and the best matching cases from the data set.
- X–ray discovery, comparison and inspection screen: to explore the uploaded image (original/pre–processed), the matching x–ray (side by side/overlay), and the other top image matches.
- Zooming and panning functions.
- Toggle between original and pre–processed images so that different previews could be viewed next to each other for comparison. This screen also allowed the practitioner to browse and select from a list of highest scoring matches.

JavaFX was used to build the user interface, enabling rearrangement or elements for rough prototyping, and detailed CSS styling of elements for the later stages of the process.

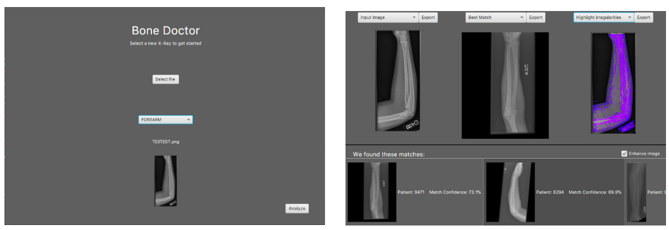

Fig 14 Screenshots of the user interface

Conclusion

Although the Bone Doctor project was not a platform well–suited to be directly used in hospitals, it was a successful proof of concept and served as an educative tool. We demonstrated a space for technological advancement in the field of radiology in common health systems. We also demonstrated that it is possible to implement a solution tackling the issue stated in the Introduction, given enough data.

Possible future improvements might include extending the dataset by collecting new user–inputted data to retrain and improve machine learning models. Or, with a dataset containing labels stating fracture type and the location of a fracture, one could, given an input image, find bones that have the same kind of fracture, to link the case under review directly to clinically similar cases thereby providing a guide to possible treatment.

Research Spotlight

How secure is your smarphone? Jonathan Goddard reports on the Digital Technology Group's study into android vulnerabilities.

Researchers in the Digital Technology group at the Computer Lab have been looking into the vulnerabilities inherent in our mobile devices.

In 2015, the team, including Alastair Beresford, Andrew Rice, Daniel Thomas and Daniel Wagner, surveyed over 20,000 devices. They used the Android app Device Analyzer, available through Google Play, which users downloaded and installed. Device Analyzer collects usage statistics and periodically uploads anonymised data for the project. The app, which was created at the Computer Laboratory, has surveyed over 30,000 Android devices to date.

The study discovered that 87% of Android devices are vulnerable, due to a lack of patching and updates from manufacturers. The vast majority of Android devices have severe security issues and did not receive patches against one or more of the 11 critical vulnerabilities that were made public in the five years leading up to the study. On average, manufacturers only issued 1.26 patches per year to customers, despite more patches being developed and available.

Most Android handsets never receive any updates to fix known problems. Often the manufacturer just moves on to the next product. Whether that's a newer handset or an app, the focus is on getting the next profitable item onto the market rather than maintaining existing products or making them more secure.

However manufacturers do vary in how secure their devices are. The results of the study placed Google first for security with a score of 5.76 out of 10, followed by LG at 4.53.

During the research study, the team developed the FUM score to compare the security provided by different manufacturers.

The score has three components:

f the proportion of devices free from known critical vulnerabilities.

u the proportion of devices updated to the most recent version.

m the number of vulnerabilities the manufacturer has not yet fixed on any device.

FUM scores out of ten:

- Google Nexus 5.76
- LG 4.53
- Motorola 3.34
- Samsung 2.81
- Sony 2.78
- Asus 2.61

By quantifying Android security issues, the team aims to provide consumers with hard data that can inform their buying choices, and make choosing a smartphone less of a minefield where security is concerned. This will in turn create an incentive for manufacturers to prioritise security for competitive advantage.

See more about this project at http://androidvulnerabilities.org/

Since this study, members of the team have gone on to investigate vulnerabilities in apps as well as in operating systems, and most recently their work has focused on third-party cookies and online tracking.

In February 2019, Dr Alastair Beresford won a Google Security and Privacy Research award.

Jonathan Goddard is Digital Communications Coordinator at the Department of Computer Science and Technology.

Computer Laboratory news

Appointments

Dr Emily Shuckburgh has been appointed as Reader in Environmental Data Science from 1st May 2019. Dr Shuckburgh is a climate scientist and has co–lead the Polar Oceans division at British Antarctic Survey, which is focused on understanding the role of the polar oceans in the global climate system

Dr Cengiz Öztireli has been appointed to a University Senior Lectureship in Computer Graphics from 1st October 2019.

Dr Öztireli is currently at the Department of Computer Science, ETH Zurich.

Awards

Professors Andrew Pitts and Marcello Fiore have been awarded the Logic in Computer Science (LICS) Test–of–Time Award 2019.

The LICS Test–of–Time Award recognizes a small number of papers from the LICS proceedings from 20 years prior that have best met the 'test of time'

Professor Pitts has also been named joint winner of the 2019 Alonzo Church Award.

The Alonzo Church Award is an annual award for Outstanding Contributions to Logic and Computation.

Outreach

Dr Noa Zilberman, Jack Lang and Dr Robert Harle delivered the Computer Science Masterclass in February 2019.

The University's Masterclass series offers taster days for Year 12 students and provides them with an opportunity to explore topics of interest beyond what is covered within the A Level syllabus, as well as the chance to experience typical undergraduate teaching at Cambridge.

Dr Zilberman delivered a lecture on cloud computing while Jack Lang gave an introduction to business studies for Computer Scientists. Dr Harle addressed the topic 'Applying to Cambridge'.

In April 2019 Dr Amanda Prorok's lab ran a workshop for 34 young women (Year 11) to gain hands–on experience with robot programming. They used 5 Thymio robots for this workshop, and the girls were tasked to code a couple of robot programs. With a little help from assistants (doctoral students), the girls successfully completed a set of exercises, creating robot behaviours ranging from simple navigation with obstacle avoidance, to having the robot dance when it drives on top of dark spots on the carpet. The workshop was a huge success, with positive feedback from the girls, and a lot of excitement about robots!

Cambridge students earn silver in national cyber security challenge

HECC (Higher Education Cyber Challenge) is a cross–university cybersecurity/ CTF competition run by the University of Southampton — following in the footsteps of the Inter–ACE challenge launched by the University of Cambridge's ACE–CSR in 2016 and hosted there for three years.

The first edition of HECC was held on Saturday 9th March 2019, at Southampton's Highfield Campus. Eleven students in three teams made the trip from Cambridge to take part, calling themselves 'cryptographic_ randomness', 'Cam Power Unlimited' and 'NP Compete'

A total of 26 teams competed, from universities across the country. Congratulations go to NP Compete (Chris Underhill, Simon Crane, Dimitrije Erdeljan and Patryk Balicki) who took 2nd place with 44700 points — narrowly beaten to the top spot by team '0x434343' from Cardiff University. After hanging around 3rd/4th position for most of the competition, NP Compete accelerated towards the end of the day and overtook team 'EmpireCTF' from Imperial College London to claim the silver medal — and a large cash prize.

The competition was jeopardy–style, with challenges covering a wide range of topics, including reverse engineering, forensics, web exploits and packet sniffing. Most were written by students and academic staff from Southampton University, with the rest provided by the competition's sponsors. Interesting challenges included an NFC problem solving task that involved scanning the lanyards of other teams to get clues for a flag, and a challenge that involved exploiting an automated VoIP call system to run shell commands.

All competitors received a free lunch and enforced breaks from hacking to rehydrate. After the competition ended, the teams had dinner, during which the results were announced and prizes given, followed by a chance to network with organisers and fellow students.

Overall, HECC was an extremely well planned and enjoyable CTF. It was an amazing experience for all of the Cambridge teams. We look forward to next year's HECC, when maybe Cambridge will take home the first prize.