CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Varghese Antony; Raja Kumar; Jonathan Lun

Project Number

S1401

Project Title

Tejon Conservancy Oak Identification App

Abstract

Objectives/Goals

We are making an app testing the accuracy and possibility of constructing it in a manner that allows newcomers to learn by going through and interacting with the tree as well as a series of options in the app to ultimately come to a conclusion as to what species they are looking at.

Methods/Materials

We took pictures which we then printed and made a physical version of the app, after that we made google forms that replicated the app. We took this info further improving on it to build it in MIT App Inventor 2. After this, we rewrote the flowchart for the app refining it and now are working towards publication.

Results

The app is actually very effective. We had an 88.8% level of accuracy amongst those that we tested on our forms. We believe that number is actually going to be higher since we have multiple versions of the app that are all improving from the previous mistakes.

Conclusions/Discussion

These results show that there are alternate routes that are available to educate people about the environment that they are surrounded by. A significant point behind this is that these people can do it on their own, self sufficiently discovering. In this sense, there is still a process of learning which keeps a sense of discovery and accomplishment. Though it is more appealing to the younger ones in this generation, the app is still available for all who want to explore.

Summary Statement

Identifying trees through an app and testing how accurate it can be.

Help Received

Scott Pipkin, Mr. Brasier

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Rohan Arora

Project Number

S1402

Project Title

Aegis: A Biosensor for Hydration Status

Abstract

Objectives/Goals

75 percent of Americans are dehydrated, leading to an influx of general health problems, such as obesity and heart disease. The reason for this is the unavailability of low-cost, user friendly hydration measurement systems. The purpose of this engineering study is to develop a biosensor for hydration that uses salivary osmolality as a biomarker, is available via smartphone, provides near real-time measurements, and exhibits at least 80 percent accuracy.

Methods/Materials

After some preliminary observations, a multidisciplinary approach was taken that combined biological theory with image processing and machine-learning techniques to create the sensor. 112 samples of saliva were collected from different individuals at different times of the day over 2 weeks. The optimal position for performing tests was determined and a case was printed to make tests rigorous for the user. Features were extracted from each vector and the Random Forest model was iterated five different times to reduce noise, and each engineered with a separate set of features. After feature selection and training the model and the system were evaluated in three different ways: OOB Error, 2-fold Cross Validation, and Cohen's Kappa (to eliminate randomness).

Results

The final model exhibited a cross validation accuracy of 87.5% and a Cohen's Kappa of 80%. It considered the average luminance of the ambient room compared to the luminance of an image of saliva, as well as the subjects's BMR. The other four iterations were substantially less accurate due to the amount of noise from external features. The final algorithm takes less than 90 seconds of processing time and requires only 4 images and basic user data.

Conclusions/Discussion

Firstly, the project verifies the usability of salivary osmolality to correspond with hydration with a larger sample set than current literature. More importantly, this technology has major implications in the field of personal healthcare, revolutionizing the ability to track a vital characteristic of health without a professional analysis. With the ability to track hydration in real time, patients will be able to maintain their general health more conveniently. An iOS application has been developed using the algorithm is accessible to everyone who owns a smartphone.

Summary Statement

This project engineered a point of care, iOS application-based biosensor with the ability to categorize an individual's hydration using personal data and images of saliva, making non-diagnostic hydration testing more accessible for patients

Help Received

None. I designed, built, and performed the experiments myself. I designed and programmed the algorithm myself after an internet search on techniques. I was simply supervised in the laboratory.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Rong Bao

Project Number
S1403

Project Title
How to Wrap a Sphere

Abstract

Objectives/Goals
My goal is to find shapes that wrap a unit sphere with small area and perimeter in order to economize material usage in wrapping spherical objects.

Methods/Materials
I defined wrapping to be a noncrossing contractive mapping of a piece of paper into Euclidean 3-space, specifically sphere here. I computed and compared four different methods of wrapping a sphere -- two circles wrapping, two circles and a strip wrapping, strip wrapping, and petal wrapping.

Results
Two circles wrapping generates a 23.35% area waste but small perimeter. Two circles and strip wrapping generates a 7.08% area waste but big perimeter. In both petal wrapping and strip wrapping, as the number of petals increases and the width of strip decreases, the area of wrapping paper decreases and will approach infinitely close to the surface area of the sphere. However, the perimeters of strip increase much more rapidly than those of petals.

Conclusions/Discussion
From my project it is concluded that petal wrapping is the most optimal method in order to minimize material usage in wrapping spherical objects. The result is also useful to problems involving unfolding spheres from 3-dimension to 2-dimension, such as map projection and disassembly and reconstruction of spherical objects.

Summary Statement
I devised and compared four different methods of wrapping a unit sphere with a flat piece of paper.

Help Received
I designed and performed this research by myself, did some computations by calculator TI-nspire, and consulted several Internet sources.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Guadalupe Bernal

Project Number

S1404

Project Title

Autonomous Robot Navigation Using Computer Vision for Exhaustive Path-Finding

Abstract

Objectives/Goals

The objective of this project is to develop a computer vision algorithm for solving the problem of exhaustively searching for a predefined path, by solely analyzing a video stream from a camera mounted on a differential drive robot. The algorithm should find all possible routes a robot can take with no prior knowledge of its environment, other than the color of the paths to be traversed. In this project I also designed and built a differential robot to test the algorithm. The path is represented by a connected graph and marked on the floor and given a completely new path the robot should still find all the routes.

Methods/Materials

The system consists of a differential robot and camera mounted on top which broadcasts a video stream to a computer through a Wi-Fi network. The computer then analyses the video and sends commands back to the robot via a Bluetooth serial modem. The algorithm was developed in Visual Studio C++ using the computer vision library OpenCV 3.1. The algorithm first uses a preprocessing pipeline to perform color segmentation and then looks for the center of mass of certain regions in a frame. After that it uses a PID algorithm to track the location of the path and determine the position of the next target. I designed the robot on SolidWorks and 3D printed the parts. The camera is an android phone which is mounted on the robot#s body. I used an Arduino DUE connected to a motor controller and two DC motors.

Results

I developed two sets of patterns which can be used to generate any possible path and to demonstrate the feasibility of the project. Experiments were conducted by recording a video of the robot following a certain pattern from above so that the top view gives a clear position of the robot as it moves through each path. Then the recorded video was replayed to measure the deviation of the robot from the path in equally divided periods of time. I limited the angle of rotation to a maximum of 30° and an equivalent radius of 20cm. The result was that robot was able to successfully complete every tested path.

Conclusions/Discussion

I developed a robust algorithm that allows a robot to follow different paths without having any prior knowledge of its surroundings and is also capable of exhaustively searching for every connected route. This algorithm is suitable for real-time processing and has industrial applications.

Summary Statement

In my project I developed a computer vision algorithm to recognize paths and navigate them autonomously.

Help Received

I took engineering courses at my high school that allowed me to learn programming and 3D computer modeling. I have also been a part of multiple robotics teams including the FTC competition where I was the main developer. My main source of information came from the internet.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Daniel Bolya; Dylan McLeod

Project Number
S1405

Project Title
Using Artificial Intelligence Systems for Autonomous Visual Comprehension and Handwriting Generation

Abstract

Objectives/Goals
Our object is to see if multiple neural networks can be used in a pipeline to parse various types of offline math problems and generate a handwritten answer as well as a human would (with at least 95% accuracy). Note: offline here means as an array of pixels, while online means as a collection of points and strokes.

Methods/Materials
Laptop and C++ compiler suite along with libraries for image processing and neural networks. We also used a lot of different openly published training data sets. The pipeline process an image of a worksheet taken from a webcam in a series of five steps: division, identification, parsing, solving, and finally writing the answer.

Results
We found the accuracy of our pipeline to be difficult to measure, as even though all the components are above 90% accurate on tests sets, the actual input of the pipeline is a real-world image which can vary greatly in quality, readability, and style. Thus, further testing is required to analyze its specific accuracy.

Conclusions/Discussion
While the scope and breadth of our pipeline does not greatly improve on previous equation recognizers#especially online ones#our applications are not deeply rooted in equation parsing, and thus many of the same concepts can be used to solve other difficult tasks such as grading simple free-response tests.

Summary Statement
We showed that a pipeline of image processing and neural networks is a good way to solve handwritten math problems.

Help Received
None. We designed, implemented, and constructed this project ourselves.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Vivian R. Chiang

Project Number

S1406

Project Title

Motion Detection in Video Surveillance

Abstract

Objectives/Goals

Video surveillance cameras are not effective if there is no one to monitor them all the time. Therefore, it would be greatly beneficial if the cameras or computers can automatically detect if there is someone walking by or if there are important motion objects to capture. I wanted to construct an algorithm that most accurately and efficiently improves current methods of separating foreground from background in video frames. My first priority is to create an original algorithm that results in a higher F-measure across all categories. F-measure incorporates both false positive and false negative counts, and is a more complete measure of accuracy of the algorithm.

Methods/Materials

I used IEEE Change Detection Workshop 2012 dataset which includes videos from the following six categories: Baseline, Dynamic Background, Camera Jitter, Intermittent Object Motion, Shadow, and Thermal. The dataset included frame by frame input images of the video and ground truth images that indicated actual motion objects. I implemented my algorithm in MATLAB. My program would detect motion objects in input images based on my algorithm, and generate output images indicating detected motion objects. My program would then compare my output images with the ground truth images to calculate true and false detection rates.

Results

My average F-measure score was 0.7149 over 31 video streams in six categories. This F-measure score was higher than the third place paper among 19 papers submitted to the 2012 IEEE Change Detection Workshop, all using the same dataset.

Conclusions/Discussion

My algorithm has four major components: camera shake detection, block histogram detection, blob cluster filtering, and activity index screening. The combination of these components performed very well with images in the Baseline, Dynamic Background, and Camera Jitter categories. The images in the Intermittent Object Motion category did not perform as well because my current algorithm lacks the function to detect motion objects that stopped moving for a long period of time. All the sets in the Shadow category also did not perform as well because my current algorithm cannot separate actual motion objects from their moving shadows, which should be considered as background. However, the overall performance of my algorithm was very good compared to published algorithms to date.

Summary Statement

I developed an algorithm to detect motion objects in videos with various types of real-world conditions for video surveillance applications.

Help Received

I designed and programmed the algorithm myself after studying prior techniques in the field.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Sanath K. Devalapurkar

Project Number

S1407

Project Title

The Algebra and Geometry of Quasicategories

Abstract

Objectives/Goals

The goal of my research was to study the quasi-2-categorical analogues of 1) abelian groups and spectra, which would generalize the notion of addition, and 2) operads and quasioperads, which would generalize the notion of a multiplication on a ring or a ring spectrum.

Results

I determined that there exists a consistent notion of stable quasi-2-categories which generalizes the notion of stable quasicategories that allows one to do basic homotopy-theoretic constructions with (symmetric monoidal) stable quasi-categories themselves. I also found that there exists a consistent notion of quasi-2-operads obtained by generalizing the notion of a fibration to the context of scaled simplicial sets, which encodes, like ordinary operads, the notion of a multiplication, but for objects of quasi-2-categories, and not just ordinary categories. These allow one to define generalizations of ordinary (commutative) rings and E_k-rings to the quasi-2-categorical world; the resulting theory can be applied to study an analogue of chromatic homotopy theory for stable quasicategories.

Conclusions/Discussion

This project sets the precedent for motivating the study of stable quasi-n-categories and quasi-n-categories for values of n>2, the foundation for which depends on the proof of an open problem in homotopy theory. More importantly, however, my research on "derived chromatic homotopy theory" and the notion of the derived moduli stack of elliptic curves could allow for a form of "derived tmf", which could potentially be extended not just to the n=2 case. This might allow for a deeper study of the relationship between homotopy theory and number theory.

Summary Statement

My research allowed for a generalization of the theory of commutative algebra using the notions of stable quasi-2-categories and quasi-2-operads which I developed, allowing me to describe a geometric approach to studying homotopy theory.

Help Received

Professor Marcy Robertson helped me a lot in teaching me about models of quasi-operads. The discussions in the homotopy theory chat room taught me a lot of interesting mathematics, which helped enrich the background knowledge for my project.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Aditya A. Dhar

Project Number

S1408

Project Title

Predicting and Identifying Forced Human Displacement Hotspots Using Multiple Linear Regression and Neural Networks

Abstract

Objectives/Goals

With almost forty million people displaced worldwide, the global refugee crisis has continued to worsen. This situation is further exacerbated by the lack of a coherent system to project the influx of refugee populations to better prepare Non-Governmental Organizations and governments worldwide in responding to the human displacement crisis. This project created a comprehensive model to analyze the factors determining the risk of forced displacement for a population, and based on the determining factors, predict the extent of displacement, allowing governments and NGOs to anticipate future forced displacement.

Methods/Materials

I used both Multiple Linear Regression analysis and Radial Basis Function neural networks with back propagation to analyze how different factors present in a country affected displacement. I gathered data from several governmental and non governmental sources for my project, and examined both persistent (economic, political, and social status of a country) and precipitating (natural disasters and conflict) factors. Using data from these sources, I developed a mathematical model, using regression analysis software, to predict displacement. I then reinforced the results using an RBF neural network, which I coded in C++ on the basis of a modified Gaussian function.

Results

The regression analysis was able to predict displacement in accuracy in over 80% of the cases, and the RBF neural network predicted displacement with the same level of accuracy as the MLR model. These models also highlighted the impact that each factor played in displacement, and supported the relationships between the factors and displacement that I had theorized (i.e. conflict was a significant indicator of displacement, and countries with worse economies had higher displacement).

Conclusions/Discussion

An important part of this project was simply being able to model global displacement. In a field where research has been reactive to displacement, applying both regression and neural networks provides the ability to be proactive, as organizations can easily predict the risk of displacement in a country. This project is unique in its ability to predict global displacement with high accuracy, its large scale coverage of over 150 nations over a five-year period, and its use of two separate modeling techniques, previously unused in analysis of forced migration, to create and validate a displacement prediction model.

Summary Statement

I devised a high-accuracy model for predicting the risk of human displacement in countries across the world.

Help Received

I designed and completed the algorithms in the project myself. I used data from NGOs and UN sources as the data for my algorithms.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Rishab Gargeya; Manan Shah

Project Number

S1409

Project Title

Automated Diagnosis of Diabetic Retinopathy in Fundus Images: A Novel Synthesis of Biological and Data Driven Approaches

Abstract

Objectives/Goals

Diabetic retinopathy (DR) is one of the leading causes of preventable blindness globally. Today, the disease has the potential to affect 382 million diabetic patients worldwide. Current screening for DR is based on manual grading of severity, leading to undiagnosed and untreated disease where screening specialists are not easily available. Less than half of DR patients are aware of their condition; the disease remains asymptomatic until later stages. The goal of this work is to develop a computer-aided algorithm for the automated gradation and detection of retinopathy severity extracted from color retinal scans. This automated algorithm may be a cost-free tool for early detection of DR, decreasing the workload on retinal specialists and allowing any diabetic patient to obtain a timely and accurate diagnosis.

Methods/Materials

We trained and tested our model on 88,702 fundus images in which disease acuteness was assessed by a retina specialist on a scale of 0 through 4. Image quality within the dataset ranged from well detailed to very noisy. Our model utilized the Python and MATLAB programming languages for image analysis and processing. Our methods for data-driven analysis include convolutional neural networks (CNNs) and intensive regression models. Biological feature extraction methods include mathematical morphology and Gabor filter banks.

Results

The quadratic weighted kappa between predicted and actual classes was 0.75 and the model's AUROC was 0.92, indicating excellent agreement between predictions and the gold standard. Our model performed as well as the average inter-rater agreement between optometrists and ophthalmologists in clinical settings (who report an average 0.75 kappa value), indicating its widespread viability. Furthermore, 80% of misclassified instances were within one class of their true value, representing, on average, small deviation from expected results.

Conclusions/Discussion

We present the first ever data-driven DR detection model based on automated feature learning. We generate applicable features in real time through deep learning with CNNs and utilize manually engineered features representing properties of prognostic biological structures. Our algorithm accounts for varying image quality and ethnicities, indicating robustness and widespread clinical applicability. We hope to extend our model to the diagnosis of other retinal pathologies and deploy an application for public use.

Summary Statement

We developed a novel algorithm that accurately diagnoses diabetic retinopathy from a simple retinal scan, using morphological image processing and data-driven feature engineering to obtain ground-breaking results on large, varied datasets.

Help Received

Our project was conducted at Stanford University's Department of Radiology under the supervision and guidance of Dr. Daniel Rubin, Dr. Luis de Sisternes, and Dr. Theodore Leng.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Dhruba J. Ghosh

Project Number
S1410

Project Title
Viseme Analysis: Implementing Contour Sequence Classification to Augment Speech Recognition

Abstract

Objectives/Goals
To test the efficiency of various lip contour detection methods as well as machine learning algorithms in using a sequence of images of a speaker to determine what the speaker was saying.

Methods/Materials
Laptop computer with Python development kit, open source computer vision and scientific computing libraries, and dataset of videos obtained from public LiLiR dataset. Implemented feature extraction algorithms and sequence classification algorithms by merging existing models and tested accuracy by running on a separate dataset.

Results
The final design, which linked adaptive thresholding, a support vector machine, and a hidden Markov model, predicted a spoken letter based on solely video data. It attained a 43% average recognition rate on the test data.

Conclusions/Discussion
I created a segmented model that took raw image data from a video and predicted the letter that was spoken. This method can be extended to cover dictionaries larger than the English alphabet, and due to its segmented nature, the extracted features can also be added to an audio-based speech recognition system. Though viseme analysis using lip contour detection and hidden Markov models cannot function professionally as a standalone program, it can be effective in filtering out noise in existing speech recognition systems.

Summary Statement
I devised an algorithm to implement speech recognition through lip reading.

Help Received
All code was written by me using open source libraries. Publicly available datasets were downloaded from the University of Surrey website.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Tanisha M. Joshi

Project Number

S1411

Project Title

A Novel Cognitive Knowledge Harvesting Approach to Train a Self Learning System for Drug Predictive Models

Abstract

Objectives/Goals

The goal is to develop an incremental self-learning model that helps in the prediction of drug behavior effectively increasing the net knowledge without human intervention

Methods/Materials

I built an iterative learning model that makes measurable progress with each iteration. I designed a BioEntity based feature-vector abstraction for capturing multidimensional harvested interactions. It taps into ontological specificity across knowledge networks to exploit existing knowledge, and explore new knowledge.

Results

My knowledge network generated unprecedented novel cross validations that include 1.16 million BioEntities, 2.008 million relationships, and 6.578m properties. The final knowledge sets were generated as a result of successive approximations of incremental knowledge gain/loss and the number of iterations have a finite upper bound that converges to an order of O(N*LogN) in terms of cost complexity. Through each iteration, new knowledge insights were generated that asymptotically and monotonically converge to the expected knowledge set with an 81% accuracy level. The P-value of the confidence aggregate score of the interactions consensus across 30 sample runs was 0.9236 at 0.05 significance level.

Conclusions/Discussion

A set of Bio-Entities is the initial input which represents a specific assay and constitutes the research problem of interest. This assay, when studied for harvesting knowledge generation created a meaningful knowledge set that more closely aligns with an associated confidence score. The generated knowledge was corroborated with a Lapatinib use case subgraph that included EGFR/HER2 receptor binding, MAPK signaling pathways reported in literature.The project successfully demonstrates an operational self learning system that can evolve, update knowledge continuously, and improve drug behavior predictions in randomized cell# based assays. Knowledge was corroborated with a Lapatinib use case sub graph that included EGFR/HER2 receptor binding, MAPK signaling pathways reported in literature.

Summary Statement

Developing iterative self-learning knowledge network topologies from curated life science datasets which converge to expected knowledge, and performing ontological feature based analysis using this knowledge to help predict drug behavior.

Help Received

Archana Gangakhedka, Drug Researcher & Dr. Lato provided verification of experimental methodology, validation of results, tips for presentation. My teacher Mrs. Taylor sponsored my project.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Adiyan Kaul; Sohan Vichare

Project Number

S1412

Project Title

Hawkeye: Unmanned Search and Rescue Missions through Intelligent Drones Guided by Computer Vision & Dynamic Pathfinding

Abstract

Objectives/Goals

Build and program an autonomous drone that can carry out search and rescue missions in realistically dangerous terrain, without human control. More specifically, the drone should be able to: 1) Fly a flight pattern around an area of land, 2) Search for and identify people, 3) Bring these people back to pre-designated safe locations, and 4) Detect and navigate around obstacles throughout the process.

Methods/Materials

Software Components: OpenCV Computer Vision Library (python), modified facial recognition classifier algorithm with 500 additional positives and 400 additional negatives for drone capability. Paper written by Sven Koenig and Maxim Likhachev (http://robotics.cs.tamu.edu/dshell/cs625/aaai02b.pdf) detailing D*Lite Pathfinding Algorithm. Computer with Python 2.6 Suite and stock math libraries.
Hardware Components: 3DR Y6 Drone Body. SF and SFP Propellers. Raspberry Pi 2 B+ Running Debian OS. 5500 mAH LiPo battery. Various wires. Soldering materials. Wi-Fi Module for Ad-Hoc Network.

Results

Drone modified by us to hold a Raspberry Pi and Camera successfully searched for and identified 2 out of 2 people in windy environment using computer vision algorithms/classifier we trained. Drone proceeded to lead said people back to safe locations while detecting/avoiding obstacles using the D*Lite pathfinding algorithm.
Hardware: Successful interfacing between Raspberry Pi Camera, Raspberry Pi, Pixhawk Drone CPU, and Drone.
Software: 1) Successful person recognition (modified OpenCV) 2) Successful spoofing of MavLink commands to control drone 3) Successful obstacle detection using OpenCV 4) Successful obstacle avoidance using D* Lite Pathfinding Algorithm (same algorithm was used on Mars Rovers Spirit and Opportunity)

Conclusions/Discussion

In the status quo, search and rescue missions remain immensely dependent on a large human volunteer base - which can be problematic in rural or dangerous locations. Our project allows for a $40 modification to an existing industrial drone that can automate it to carry out search and rescue missions independently - something which we believe will be immensely useful to government organizations. Secondly, we have developed a modular way for a Raspberry Pi and Camera to interface with a common drone CPU (the Pixhawk) and create an Ad-Hoc network for computer connection - something that any developer can use to automate drones.

Summary Statement

We created a $40 method of automating industrial drones to carry out unmanned search and rescue missions in realistically rough and windy terrain.

Help Received

None, save for open source libraries (OpenCV, math libraries) and this paper: Pathfinding with D* Lite (http://robotics.cs.tamu.edu/dshell/cs625/aaai02b.pdf)

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Sarah H. Kazmie

Project Number

S1413

Project Title

DermatoScan 2.0: Image Analysis Optimizations for the Early Detection of Skin Cancer

Abstract

Objectives/Goals

To develop a simple, responsive and reliable method of digitally assessing the risk of cancer in skin lesions using an iPhone app.

Methods/Materials

8 numerical methods were defined and implemented. An evolutionary machine learning algorithm was developed to optimize the weights of each factor in the analysis. A#Complexity# method was then developed and added to the original analysis network, increasing the sensitivity and specificity of the resulting assessments. Each feature extraction method was optimized and tested, yielding a 300% average acceleration with no loss of reliability. The methods were ported from C++ into Swift. A new class was created which incorporates many of the feature analysis processes and replaces the image manipulation functionality originally provided by OpenCV.

The app tested successfully on preloaded images, but analysis performance on an iPhone was extremely slow. A new user-interface was designed to allow the user to adjust three interactive parameters, enabling the user to limit the size, color-space and border isolation threshold of the captured image. When the iPhone camera failed to deliver reliable images, the app was tested against enlarged image samples displayed on a second computer screen.

Results

The analysis process consistently achieved better than 98% accuracy within its test sample image set. The embedded DermatoScan iPhone app achieved similar results when the iPhone camera captured enlarged images but could not achieve consistent results capturing actual size samples with the iPhone#s camera.

Conclusions/Discussion

This experiment was conducted over the course of a year, building on the work of my previous experiments, based on research over the previous year. The success of this demonstration supports two of the three components of the hypothesis. The code produced consistent and reliable results even after substantial optimization. Optimization efforts achieved a 3:1 average acceleration from the improved feature extraction methods and an additional 10:1 average acceleration from the user-guided optimizations.

The code was successfully translated and embedded in an iPhone app. The iPhone camera had difficulty focusing and the internal flash proved ineffective at such close range, but these issues might be mitigated in the future with additional image processing, an optical lens attachment or an alternative external camera device.

Summary Statement

I designed, developed and tested an iPhone app to record, analyze and assess the risk of malignancy in skin lesions.

Help Received

Help Received: My dermatologist initially explained the characteristics which commonly indicate malignancy. I showed DermatoScan to 2 engineers and a scientist at Amgen, who suggested a possible attachment to achieve a more consistent image with the iPhone camera. A family friend introduced me to

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Preetha S. Krishnamurthy

Project Number

S1414

Project Title

Compression Based Data Mining for Assessing Similarity in Object Textures in Multiple Images

Abstract

Objectives/Goals

The goal of this project is to analyze if a given image is similar to any of the images in a large dataset, and if so, how similar. While comparing two images, we seek to also account for distinctions in texture in the objects captured in the images. The approach must be simple (low complexity) and should not require extensive training (as in machine learning).

Methods/Materials

The code was written and executed using Matlab and its tool suite, on a Lenovo laptop computer. The images considered were from datasets of (i) toolmarks on metal obtained from the Internet and (ii) a dataset that we created by cutting up two types of wood into blocks and taking pictures.

Results

The results from this project demonstrate that the CK algorithm can quickly analyze a large image dataset and pick out those images that are most similar (best matches) to a target image with an accuracy of 100%. It is resistant to distortions in the image; specifically it can handle cropping, erasures, and noise/speckles and yields an accuracy of between 80 to 100 % with such distortions. It significantly outperforms an algorithm that uses the Euclidean distance measure as a similarity assessment to compare two images.

Conclusions/Discussion

In this project, we seek to tackle the hard problem of examining if an image is similar to any of the images in a large data set via an automated technique. The motivating application is criminal forensics. We perform a clever application of what is called the Campana-Keogh (CK) algorithm to quickly, and yet accurately obtain the aforementioned similarity assessments. The algorithm exploits a key property of MPEG encoding. Specifically, if the images to be compared are concatenated and encoded as a video, similar images will lead to smaller encoded sizes. We conduct extensive experiments to evaluate our approach with two datasets (a metal dataset that we obtained from the Internet, and a wood dataset that we developed). We show that the algorithm provides accuracies of 80 % to a 100% with both datasets even when the images are subject to various forms of distortion.

Summary Statement

My project seeks to compare a given image with a large set of images in a dataset with high accuracy but low complexity, using compression based data mining.

Help Received

I recieved guidance from Professor Eamonn Keogh at UCR, who taught me the concepts behind the CK algorithm as well as video compression. He also provided guidance on how to develop the code. The supporting algorithms were built by me.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Rajiv Movva

Project Number

S1415

Project Title

Halo: A Machine Learning Assisted System to Detect Concussions in Sports

Abstract

Objectives/Goals

Every single fall, collision, or strike puts an athlete#s brain at risk. This danger manifests as 4 million sports-related concussions occurring annually, of which an estimated half remain undiagnosed. Detecting concussions immediately is critical to avoiding future damage that often leads to permanent neurological disability. Thus, the goal of my system is a real-time concussion detection system that is accessible to athletes at all levels of play.

Methods/Materials

A biomechanical model of the head was designed in ScanIP by concatenating MRI data (from the Alzheimer's Disease Neuroimaging Initiative) and meshing the 3D volume to a model containing ~250,000 elements. To quantify concussive tissue damage, ten NFL impacts were reconstructed by performing finite element analysis on the model in COMSOL Multiphysics. An artificial neural network was trained to classify accelerations as concussive using 358 impacts simulated in the finite element model. The system hardware layout was designed in DipTrace, and PCB production was outsourced.

Results

The neural network was validated with 57 NFL impacts reconstructed in the finite element model after video analysis. 53 of 57 impacts were correctly identified (accuracy: 93%), with sensitivity 98%. The current revision of the system costs ~$20.

Conclusions/Discussion

Here, the neural network allowed for approximation of the complex computational processes modeled in the finite element solver. Thus, the cost of realtime concussion detection was reduced to an affordable amount while maintaining a high sensitivity. This system will aid in prompt identification of brain injury, significantly reducing athletic risks the world over.

Summary Statement

Halo is an affordable system that can be worn by athletes to detect sports concussions in realtime.

Help Received

Dr. Nelson helped explain some concepts in physics and reviewed a draft of my technical paper. My brother gave me tips on poster design, and my parents provided support.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Mihika Nadig

Project Number

S1416

Project Title

Application of a Deep Learning Architecture Using Convolutional and Recurrent LSTM Networks to Video Classification

Abstract

Objectives/Goals

Currently, the use of neural networks has been restricted to image classification. Video classification has been a challenge due to the difficulty of fusing spatial-temporal information. My project aims to construct a deep learning architecture involving convolutional neural networks (CNNs) and recurrent neural networks (RNNs) with LSTMs as a viable solution for video classification.

Methods/Materials

To perform feature computation, two CNNs are used to process raw individual frames and optical flow images from the UCF-101 Video, CIFAR-10, and Microsoft COCO Datasets. After filter learning, feature aggregation is performed via feature pooling or RNNs with a deep LSTM architecture. Three feature pooling architectures were experimented with: Conv Pooling, Intermediate Pooling, and Neighborhood Pooling. A softmax classifier, the final layer in the feature pooling architecture, returns the prediction. Training is done by coupling stochastic gradient with momentum at 0.9 for optimization processes, and weight decay of 0.0005 was used with a learning rate of N frames ∗ 10−5. The LSTM Architecture provides output by returning the prediction at the last time step.

Results

Conv Pooling had the best results at 89.6% when using the three datasets. Comparing the Conv Pooling with the deep RNN using LSTMs, the second feature aggregation method, LSTMs outperformed Conv Pooling at 90.5%. These are comparable results for modern usage and demonstrate that identifying long range temporal relationships between video frames is crucial to video classification.

Conclusions/Discussion

I was able to surpass the benchmark of 2% by over 65% instead of using the naive method of performing 3D convolutions over each single frame. Because Conv Pooling was optimal, it was concluded that max-pooling over the outputs of the final convolutional layer is important. It is clear that LSTMs performed better due to its ability to identity long range temporal relationships. By coupling optical flow and raw image (300 frames/video) as input, I improved previous work that only sampled 60-120 frames.

Summary Statement

I created a deep learning architecture coupling CNNs and RNNs with LSTMs to aid in the problem of video classification, which can be extended to applications ranging from military assistance to navigational technology for the blind.

Help Received

I would like to thank my dad for providing support and resources for my project.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Gabriel Nahum; Zachary Neiman

Project Number

S1417

Project Title

C-Color: Colorblind Smartphone Application

Abstract

Objectives/Goals

Over 560 million people are colorblind. These people cannot see all colors and will have a hard time distinguishing between two colors that, to them, are too similar to each other. For example, they have problems when they can#t see where writing appears on a similar colored background. The application C-Color solves these problems by being a mobile application for Android and iOS devices. The application labels the color of the selected part of the screen to help the user know what color it is. The app also has a mode where the user can capture an image and press checkboxes to change certain colors to either black or white. This color changing mode can help the users distinguish between two similar-looking colors by making one of those similar-appearing colors dramatically different.

Methods/Materials

At the beginning of our design process, we looked at customer surveys to see what was great about our initial prototype and what needed fixing. Also, we tested the ability of the app to identify colors to see if the labeling feature needed adjustment. The app was tested on 36 different colors (3 times each) in normal, dark and bright environments to see how many times each color was identified correctly. These environments were created by darkening and brightening the image of the tested color range. There is also a customer feedback survey for our current app with pending results. These results will be used to update the app even further.

It costs $99 a year to be able to publish apps on the apple store. Google play has a $25 one time fee for a developer#s account. Android Studios, to make the android app, and Xcode, to make the iOS app, are free. This brings the total cost to $124.

Results

The original survey tests were very positive about the concept of the app. The color-identifying tests showed that most colors were identified correctly. The brightened test was a bit worse in sensing the correct colors than the dark and normal tests, but the wrongly identified colors were still very close to the actual colors.

Conclusions/Discussion

These tests don't only test the labeling function, but also test the color-changing algorithm. Because the colors were off by very little in the tests, the application is very successful because people have slightly different limits of what they consider one color or another. For example, one person may see pink when another person and C-Color may see purple.

Summary Statement

We created an effective application that helps colorblind individuals distinguish between colors and understand what colors are present.

Help Received

We designed and built the application and ran the surveys and tests ourselves.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Kevin K. Pho

Project Number

S1418

Project Title

Digital Fingerprints: Constructing One-Way Hash Functions

Abstract

Objectives/Goals

The objective of this experiment is to test numerous designs (based on mathematical problems and basic operations) to aid in the construction of faster and more secure hash algorithms. The designs were evaluated by speed and the lack of collisions, which occur when two inputs share the same output. The construction of more efficient hash algorithms provide security of passwords, integrity of files, et cetera by preventing inversion of the function and the discovery of a collision.

Methods/Materials

A laptop computer with the Python interpreter was used to design and test the algorithms. The computer and interpreter provided the timer for testing speed. The algorithms were tested on a sample word list of common passwords (to demonstrate applicability). Some of the algorithms were derived from source code or programming libraries, such as hashlib (Python).

Results

The data indicates that primitive operations performed quickly, clocking in at under a second. They provided many collisions (most yielded more than 500). The fastest operations included addition, XOR, and bit-shifting while the most secure included bit-shifting, addition, modulus, and subtracting. On the other hand, the investigated mathematical constructs that were the fastest included SHA-3, MD5, the position-based design, the addition construct, and the linear congruential generator while those that provided the most resistance to collisions included SHA-3, MD5, the sum of three cubes design, the linear shift feedback register, and the discrete logarithm problem design.

Conclusions/Discussion

The speed of primitive operations allows hash algorithms to quickly perform operations on the bits in order to "mix" them. Although the quantity of these operations affects the speed of the algorithm, the security of these do not create most of the security of the algorithm; the mathematical constructs do. The constructs are the steps that properly transform/diffuse the bits, so that the output provides a seemingly random output. Furthermore, if the constructs are based on computational hardness assumptions, they can be further proven to be secure as an input must be impossible to retrieve from its output. This information encourages the construction of hash algorithms with NP problems (problems known to be hard to solve in polynomial time), and it can aid in the construction of future hash algorithms to further enhance cryptographic security.

Summary Statement

A hash function is most effective (measured by speed and lack of collisions) when simple mathematical problems, such as the discrete logaritihm problem, are used in junction with primitive operations, such as addition and XOR.

Help Received

I designed and programmed the hash algorithm and evaluating tests and analyzed the data myself. Some of the algorithm designs were based on pre-existing ones, such as MD5, SHA-3, and Adi Shamir's Discrete Logarithm Hash Function.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Kalyani Ramadurgam

Project Number

S1419

Project Title

Improved Machine Perception During Autonomous Driving via Semisupervised Deep Learning and Multisensor Datastream Fusion

Abstract

Objectives/Goals

The objective of this project was to build a machine perception system capable of fusing sensor data from disparate sources to give an autonomous car situational awareness. Traditional computer vision systems perform poorly with multiple sensor streams and 3D reconstructions because relatively little work has been done in the field of sensory fusion, especially with regards to extracting information from the overlap of feeds.

Methods/Materials

This project proposed a method of tensor stacking, processing a stream of fused sensory input in conjunction with a new method of generating initial weights in a semi-supervised custom convolutional neural network. This architecture generated semantic identification and tracking tags of objects in the car's surroundings in real-time. Google's TensorFlow library was utilized in the execution of some elements of the cnn.

Results

The proposed model was tested with the KITTI benchmark dataset and compared to scores generated with individual sensor feed processing in traditional convolutional neural networks, the current standard in autonomous vehicle detection. The accuracy of pedestrian detection saw an improvement from 64.70% to 79.01%, and the oblique cyclist detection accuracy increased from 61.37% to 68.92%. The semi-supervised method of weight initialization proved to support the speed of the algorithm as data size increased.

Conclusions/Discussion

The results validate the efficacy of such tensor stacking and sensory fusion routes and faster detection of pedestrians and cyclist at the most oblique orientation. Additionally, semi-supervised weight initialization can be implemented in many other applications of neural networks, increasing the speed of each processing layer as the network learns. As a result, the proposed methods show potential in improving the situational awareness of autonomous vehicles.

Summary Statement

Extracting information from overlapping sensory data feeds and a new method of weight initialization to make autonomous vehicles more robust.

Help Received

My brother offered guidance in navigating and implementing Google's TensorFlow.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Tejas N. Rao

Project Number

S1420

Project Title

Reduce Accidents with Automobile Driving Tracker Using Raspberry Pi, Python, and OBD Port

Abstract

Objectives/Goals

The objective of this study is to determine over speeding, fast braking and other driving hazards to reduce accidents and increase fuel efficiency. By capturing data from automobile engine via an On Board Diagnostic Port and identifying risky driving we can track driver behavior and provide feedback for better driving habits. This is especially helpful for teen drivers, for whom accidents are #1 cause of death.

Methods/Materials

Connect car OBD port to Raspberry pie with Bluetooth protocol. Write a Python program on Raspberry pie to read engine data and log to file. Add GPS location to log. Collect logs for each trip. Upload logs to Google Cloud to analyze and highlight risky driving. OBD port Bluetooth reader, Raspberry Pi, Python Programs and Google Cloud.

Results

From multiple test drives car engine data was captured for speed, rpm, gps and engine temp using a Python program. GPS data was overlaid on top of speed to determine exact location of overspeeding and visually displayed on Google maps. Log files were analyzed on Google Cloud and from charts the number of times of hard braking was determined. Charts and Maps were used to provide feedback to drivers.

Conclusions/Discussion

Car/Auto/trucks can be monitored for risky driving using OBD port reader and Raspberry Pi. The data can help reduce reduce accidents and damage to life and property, specially teen deaths. The solution can also help with diagnosing engine faults and fix them before car breaks down. Also fule efficiency can be increased by better driving habits.

Summary Statement

Reduce Accidents and save lives with Automobile Driving tracker using Raspberry Pi, Python and OnBaord Diagnostic Port

Help Received

I got help with Python programming.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Shalin V. Shah

Project Number

S1421

Project Title

Voice: Computer Vision Algorithms to Counter Age-Related Macular Degeneration, Glaucoma, and other Pathological Typhlosi

Abstract

Objectives/Goals

Most everyday items are not formatted for Braille and audio. Furthermore, of the 7.3 million blind people in the US, 90% can#t even read braille. So blind people are very limited in the amount of content they can read. The goal of this research was to build an algorithm that can effectively guide a blind person into taking a picture using assistive computer vision, and then read the words that are on that picture aloud. The research also improves the accuracy of the reading through the various image modifications.

Methods/Materials

The Voice algorithm was built in the programming language of Objective-C. It uses Google Drive#s Application Programming Interface for the Optical Character Recognition because of the multi-language functionality. Also, the front end of the application uses the iOS platform and thus requires either an iPhone, iPad, or iPod to run. The image processing occurs in the background. The algorithm was tested in both light and dark condition with a variety of distances ranging from 5 inches to 2 feet. Additionally, it was tested on various reading material such as newspapers, museum placards, medicine labels, expiration dates, can labels, and more. It was also tested by many blind people over a spread out time interval. Each time, new functionality was added in order to fix the problems that were exposed and improve the usability for blind people. So over multiple iterations, the application improved through added functionality.

Results

The results proved that guiding the blind user through taking a photo did assist them greatly in using the front-end application and taking good quality photos. Additionally, the silent photo manipulations done by the algorithm also help increase the quality of the spoken text because of the improved lighting, contrast, and orientation on the modified image.

Conclusions/Discussion

The algorithm can successfully guide the blind user into taking apt photos of many different items that are not formatted for braille or audio and read the words from the images out loud in a matter of seconds. It can read those words in around thirty different languages. Also, it can read many pages one after another without using additional time because as one image is being read, the next is being processed. So it can read items such as magazines and book pages one after another.

Summary Statement

Using computer vision algorithms, I created an application that guides blind people into taking a photo of everyday items (not available in Braille/audio) like nutrition facts and medicine labels, and reading the words from that image aloud

Help Received

None. I designed, built, and performed the experiments myself.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Nikhil Sundrani; Sameer Sundrani

Project Number

S1422

Project Title

The Use of Photoplethysmography (PPG) to Develop a Prototype Heart Rate Wristband to Monitor Life Threatening Arrhythmia

Abstract

Objectives/Goals

The objective of our project is to save the lives of the nearly 300,000 people that die of cardiac arrest in the United States alone, every year. These peoples deaths are preventable, as they only die because they do not make it to the hospital in time. Thus, to save these people, we created the proof of concept of a wristband that will continually monitor heart rate and alert EMS from the phones of those in danger of imminent cardiac arrest.

Methods/Materials

We used an Arduino microcontroller, which is a prototyping shield, and its associated coding platform. Using these, we were able to connect a PPG (photoplethysmography) sensor, and use this advanced technology to monitor heart rate. The coding that we implemented within the platform was both source code and modifications to this code to make the heart rate algorithm, and and achieve our goal. We tested it ourselves for bradycardia, tachycardia, and sinus rhythm (heart arrhythmias), which it successfully recognized. The phone application was developed using Blynk software (and IOS and Android application) which is customizable as per the creator#s preference.

Results

The PPG sensor system prototype read various arrhythmias with maximum accuracy, producing results and notifications as soon as the given arrhythmia was detected. Notifications regarding EMS were received via the phone application. Given the different parameters between 40 beats per minute (BPM) and 200 BPM, the notification was triggered 100% of the time, with 100% accuracy.

Conclusions/Discussion

We created a proof of concept prototype PPG sensor system using the Arduino platforms and hardware that allows for the detection of sinus rhythm and the other arrhythmias that we were examining and analyzing: ventricular and supra-ventricular tachycardia, ventricular fibrillation, and severe bradycardia. Since the system that we designed notifies the user on their smartphone via the Blynk application as soon as one of the above arrhythmias are detected, it is probable that people who have life threatening arrhythmias (that lead to cardiac arrest) will be able to receive treatment from a hospital on time.

Summary Statement

We developed a prototype heart rate wristband that has the application to notify EMS and potentially save those with life threatening cardiac arrhythmias.

Help Received

We designed the system ourselves after research of the needed algorithms and had some help from our teacher, Dr. Gettman, in working out issues in the coding.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Robert G. Tacescu

Project Number

S1423

Project Title

Safecopter: Developing a Collision Avoidance System Based on an Array of Time-of-Flight 3D Cameras

Abstract

Objectives/Goals

Multicopters have a wide range of applications from surveillance to package delivery and medical support. Although growing in popularity, they are not used yet on an industrial scale for safety reasons. The goal of Safecopter is to develop a modular collision detection and avoidance system that would make flying a multicopter in autonomous or tele-operated mode completely safe. Integrating a system of sensors that provides 360° spherical coverage and works reliably in any light conditions opens the door to the safe use of multicopters and to an extended range of applications.

Methods/Materials

Based on the research of a wide variety of sensors, I reached the conclusion that the best solution would be to use a Time of Flight 3D camera array. To be able to process the point cloud produced by the 3D cameras, I use a compact onboard computer running the Ubuntu Linux operating system. The software system is programmed in C++ using ROS (Robot Operating System) as a development platform.

Results

Data received from the 3D cameras is processed and combined into a single point cloud representing all the objects around the multicopter up to six meters away. Based on the vector of travel, if a collision is detected, a new trajectory is calculated. For visualization and debugging purposes, this cloud of points is categorized and displayed in different colors on the ground station computer.

Conclusions/Discussion

Based on testing performed, the system can reliably detect possible collisions with objects in different lighting conditions (day/night, indoor/outdoor, etc.) and reroute the multicopter to the shortest path to the target.

Summary Statement

The goal of my project is to create a multicopter that doesn't collide with nearby objects. It is based on a modular collision avoidance system that includes an array of 3D cameras mounted on the top to detect and avoid objects in the way.

Help Received

None. I designed, built, and programmed the system myself.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Michelle C. Xu

Project Number
S1424

Project Title
Determining the Protein Structure from Ant Colony Optimization Using Energy Minimization Derived from the Ising Model

Abstract

Objectives/Goals
Fusion inhibitor proteins prevent the HIV-1 gp41 glycoprotein from binding with the receptors on the host cell. However, scientific understanding of inhibitory potencies among different inhibitors remains unclear. This project proposes a novel computational approach to investigate the protein folding mechanism, with focuses on the fusion inhibitor proteins just by looking at their amino acid sequences.

Methods/Materials
I developed a new protein folding energy function using the idea of atomic spins from the Ising model. In this new function, the atomic spins were extended into vectors to simulate the shape of the protein peptide bonds. Vector space algebra was then applied to transform the statistics-based Ising model into a deterministic form. I applied an artificial intelligence method, the Ant Colony Optimization with a deterministic searching algorithm, to solve the energy minimization based on the new folding energy function by looking for the global minimum in the energy graph. The computational results of selected fusion inhibitor examples were compared with nuclear magnetic resonance (NMR) and X-ray crystallography results for the same proteins from the Protein Data Bank for similarity validation. Similarity was validated through visualization using the 3D viewer Jmol, as well as numerical comparisons through a topological property called Relative Contact Order (RCO).

Results
The results showed high similarities in both visual and numerical comparisons between the calculated and the actual protein structures. In addition, the HIV-1 inhibitors with weaker binding strength had multiple minima in the folding energy curve and the global minimum was found at the misfolded conformation.

Conclusions/Discussion
My protein folding energy function allows me to examine both the native folded and misfolded conformations. I was able to find that for the fusion inhibitors with less inhibitory strength, the energy curve showed a global minimum at the misfolded conformation rather than at the native folded conformation. This discovery could improve drug design efficiency by contributing to the proper selection and modification of the HIV-1 fusion inhibitors, as well as other anti-virus drug design projects.

Summary Statement
I was able to develop a computational approach using vector space algebra to solve for the protein native structure by minimizing the energy through an artificial intelligence method called Ant Colony Optimization.

Help Received
I would like to thank Professor Stephen H. White from the UCI Department of Physiology and Biophysics for his mentorship by helping me narrow down my research topic and providing me with research articles. I would also like to thank my math teacher, Mr. Charles Y. Beilin, for his support in the development of

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Alexander M. Zlokapa

Project Number

S1425

Project Title

Predicting Future Body Mass Index with an Artificial Neural Network

Abstract

Objectives/Goals

Despite the widely known health risks of obesity, there is no accurate and reliable way to predict if an individual will become obese in the future. The goal of my research was to create an artificial neural network to predict if a person will have a normal, overweight, or obese body mass index over a decade into the future, thus creating an early warning system against obesity.

Methods/Materials

Data was obtained from the British Cohort Study of 1970. A neural network was written in pure Java without using any external libraries or code. It was then tested on the data of 1125 real people in the dataset, and its results were compared to the performance of a logistic regression and support vector machine (two traditional models from the LIBLINEAR machine learning library). The Apache Commons Math Library was used in the computations for the statistical analysis of the results.

Results

Testing for future obesity, the neural network had a conclusively high positive likelihood ratio of 14.7 (95% CI [10.3, 21.2]), while the traditional models inconclusively predicted obesity. Overall, the neural network had the greatest classification success of the three models tested, and its misclassifications were closest to the correct value. A web application was developed to apply the neural network's success in the real world, allowing users to enter their data and obtain a prediction.

Conclusions/Discussion

This research presents the first successful and conclusive prediction tool for future obesity. This is also the first known application of neural networks in adult obesity, demonstrating the superiority of neural networks over traditional models for prediction.

Summary Statement

I created an advanced artificial neural network that is the first successful and conclusive prediction tool for future adult obesity, significantly outperforming traditional models.

Help Received

None.