

SUCCESS STORY

AI Cash Recognition for the Visually Impaired



Using AI to Help Visually Impaired People Identify Cash Notes

Individuals with impaired vision now have a means to use their smartphones to confidently perform monetary transactions in the marketplace. Thanks to deep learning and members of a community project that developed the smartphone application, bank notes can be quickly identified and their value conveyed to the user through spoken audio.

“Because we are only seeing the beginning of the use of artificial intelligence (AI) to aid the visually impaired and because of the rapid pace of innovation, the use of AI will continue to become more and more streamlined with our lives. As technology continues to be adapted and tested, there is hope that life for the visually impaired will see far fewer obstacles.”¹

- The Vision of Children Foundation,
San Diego, CA

Challenge

For individuals with impairments to their vision, simple acts, such as making purchases or performing other monetary transactions with cash notes, present a challenge. This can be frustrating and can even limit social interactions.

Solution

A smartphone app relies on artificial intelligence (AI) techniques to positively identify cash note values through the built-in camera. The app then reports the values audibly to the visually impaired person.

Background and Project History

A native of Nepal, Kshitiz Rimal is deeply involved in community projects that use AI applications to improve the lives of people. As an Intel® Student Ambassador for the Intel® AI Developer Program, he launched the Cash Recognition for Visually Impaired project. Among a broad range of other AI-related endeavors, Kshitiz also serves as a City.AI* Ambassador for Kathmandu City and Fast.ai* International Fellow. A not-for-profit organization that he co-founded, [Artificial Intelligence for Development](#), researches and implements deep learning projects for tackling societal problems, trains students in the uses of AI, and works with government to identify and resolve community issues. Additionally, he is a Google Developer Expert (GDE) on machine learning.

The project initiated by Kshitiz—[Cash Recognition for Visually Impaired](#)—empowers people with vision loss to correctly identify cash notes and make purchases and transactions with them. Originally begun as a project to offer greater independence to visually impaired individuals living in Nepal, the AI techniques developed over the course of this project adapt well to many other identification tasks that prove difficult for anyone with vision loss.

The AI-powered app uses an integrated smartphone camera to capture the images of cash notes being held by a person. After recognizing the value of the note from the image using a deep learning model that is embedded in the app, an audio sequence plays that signifies the value of the note.

Kshitiz gained interest in this project through a chance encounter with someone on a bus. “I was traveling on a public bus,” he said, “and sitting next to a blind person. When his bus stop came, he asked me to tell him about the value of the monetary notes he was holding. I answered him immediately, but as he was walking

towards the bus exit door I then asked him if it's a problem identifying these note values on daily basis. He told me that he had recently gotten into an accident and became blind and ever since it was troublesome for him to identify notes by just touching them. He also shared that it's more common for people who are not blind at birth, but become blind later. This got me thinking and eventually I came up with the idea for this AI project."



Figure 1. Kshitiz Rimal is an AI enthusiast, technologist, and Intel® Student Ambassador.

An Abiding Interest in Computers

"Since my childhood," Kshitiz said, "I have been fascinated with computers and technology in general. This passion led me to where I am today."

Kshitiz noted that AI to him is a natural progression of his earlier career path. "I think AI is the greatest tool that we humans have created and because of that we need to use it for social and humanitarian benefits."

Finishing the Project

It took more than six months for Kshitiz to make a fully functioning app that could work entirely offline and recognize all cash notes in Nepalese currency. The deep learning model is now complete, as is a fully functioning Android* mobile app.

"Right now," Kshitiz said, "I am fixing some bugs and issues related to machine learning on the iOS* platform. After that, the project will be completely finished and it will work offline with cross platform support as well."

Today, approximately 200,000 individuals in Nepal are visually impaired. The Cash Recognition for the Visually Impaired project will directly help these individuals perform monetary transactions and empower them in their daily lives. "I hope this project," Kshitiz said, "will further inspire the community to develop similar solutions that can directly affect the lives of others—taking advantage of AI technology."



Figure 2. Nepalese banknotes lack clear identifying markings for those with impaired vision.

Developing the Solution

As documented in [more detail on the Intel® AI Developer Program site](#), Kshitiz overcame a succession of challenges creating and enhancing the operation of the cash recognition app.

Prototyping the Approach

The first step in this project was to capture enough images of notes in two currency categories (Rs.10 and Rs.20 notes) to perform training of the deep neural network to see if the planned approach was workable. The first batch of 200 images (see Figure 3) included different angles and viewpoints for each of the note categories.

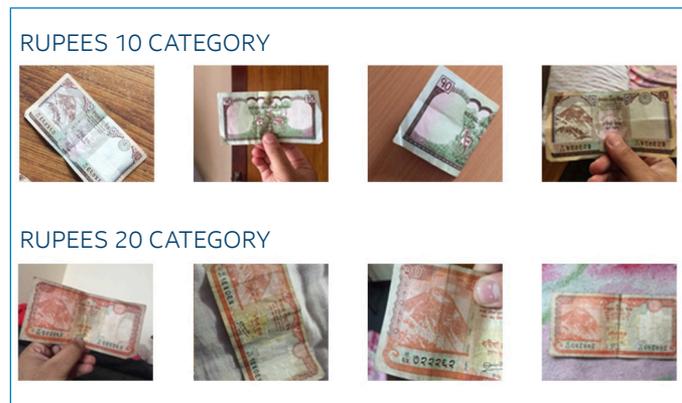


Figure 3. Multiple images of 10- and 20-rupee notes.

Transfer Learning Stage

Training a deep neural network for image recognition typically requires a large dataset of images, but by using transfer learning techniques, the recognition can be adapted to a small number of datasets. Kshitiz commented, "What we do is take a model that is already trained on a huge dataset and leverage its learned weights to re-train on the small dataset that we have. That way, we will not require a large dataset, and the model will predict correctly as well."

At this stage, Kshitiz is still confirming that his approach will be effective. He used a VGG16 model pre-trained on a popular image-based dataset with 1000 categories. By using a fine-tuning technique on VGG16, he only needed to re-train the last layer of the model (as shown in Figure 4) to achieve the right degree of accuracy.

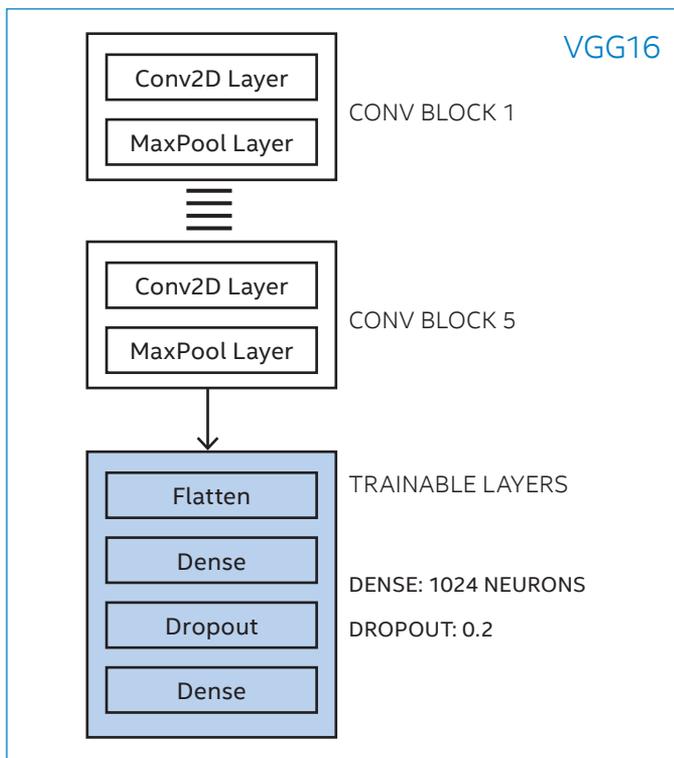


Figure 4. Retraining the last layer of the model.

The fine-tuned model had the following characteristics.

Table 1. Initial training configuration

Number of epochs	50
Total number of training data points for both categories	400
Total number of validation data points for both categories	300
Learning rate	0.001
Optimizer	RMSprop
Loss	Cross Entropy Loss

Kshitiz ran the training using Keras* with a TensorFlow* backend for the code, with a validation accuracy of 97.5 percent, whereas the training accuracy is 98.6 percent. He determined from these training results that the selected model was slightly overfitted. Additional data would be the best way to correct this issue, as well as introducing regularizations to the model.

Developing the App

To interact with the prototype being developed, Kshitiz created a minimal viable product (MVP) version of the app using React Native. Through interactions with a REST API, the two-category version of the app tapped into compute operations performed on the server to carry out recognition and prediction. The response is sent to the client app following prediction. Display of the labels and the audio output is completed on the client side by the app.

Further Refinements

The work accomplished to this point established that the approach was viable. The next stage of the project involved collecting image data for the seven categories of bank notes commonly used in Nepal. Starting from an initial value of 500 images, Kshitiz gradually increased the number to 2500 images for each category to achieve the desired level of accuracy.



Figure 5. An offline version of the cash recognition app was developed for iOS* and Android* smartphones.

Once training the model with the expanded image sets was done, Kshitiz rebuilt the app so that it could function offline without Internet access. Additional pre-training of the model type to VGG19, using 5,000 images for the two categories (2,500 for training and 800 for validation for each instance) achieved an accuracy rate of about 95 percent.

Some of the lessons learned in the course of this development project are documented an article that Kshitiz wrote and published on Medium.com: [Practical Lessons Learned While Implementing an Image Classifier](#).

“Because of the support from Intel while developing the project, I was able to complete it properly. Intel particularly helped with technical expertise I required to complete this project by connecting me with various experts in the team. Also, Intel provided me with hardware and computation resources through the Intel® AI DevCloud, access to software development kits, and access to technical learning resources to develop and train the model. They also helped me publish posters, blogs, and articles to increase exposure to my project around the world and provided me access to other developers and AI enthusiasts through the Intel® AI Ambassador program.”

- Kshitiz Rimal, Intel Student Ambassador

Enabling Technologies

Intel® AI DevCloud, providing free access to a high-performance platform for AI training and development

Intel® Distribution for Python*, accelerating application performance with minimal (or no) modifications to code

Intel® Optimization for TensorFlow*, a machine-learning framework for efficient computation used to train the model on Intel AI DevCloud

“I am planning,” Kshitiz said, “to make this app available in stores and public places, using technology such as the Intel® NUC (Next Unit of Computing) PC and Intel® Neural Computer Stick, which can accelerate AI inference on the edge for fast offline prediction and recognition tasks.”

Kshitiz found Intel AI Developer Program immensely helpful to the tasks at hand, particularly the video tutorials and guides. Intel-optimized libraries for developing and training the deep learning model were also helpful for accelerating the overall project development. “The Intel AI expert team,” Kshitiz said, “helped me a lot while refining technical details, optimizing code to run on Intel AI DevCloud, and solving common issues during development.”

“There are lots of things that can be developed based on the idea of this project,” Kshitiz said, “that can empower visually impaired individuals in their daily lives. But I see this as a community project where interested individuals can contribute to it. All the code and datasets will be made publicly available.”



“I think that based on this project’s idea, a lot of similar things can also be developed that can have commercial applications. For example, using the same approach, we can create an app that can recognize daily grocery items in the supermarket and lets visually impaired individuals know how many boxes of certain items are present in front of him or her. This can enable such individuals to navigate the market independently and confidently.”

- Kshitiz Rimal, Intel Student Ambassador

Figure 6. Indian street seller in Kathmandu, Nepal.

Tips for Other Developers

Reflecting on the work accomplished so far, Kshitiz noted, “To ideate a project is one thing but to complete it properly is a different story. While working on this project, I got exposed to the end-to-end building process of how to develop a solution that utilizes deep learning technology. Throughout this time, I got a sense of what is required to develop a solution that is useful in the real world. No matter how small the project is, or how simple your idea is, when you start to build it from scratch, you will learn so much and gain experience that you won’t get from anywhere else.”

The few issues that Kshitiz struggled with during the project were primarily related to building the actual model, although some involved developing the application code for the mobile app. He advised that, particularly while leveraging transfer learning techniques, it is essential to determine how many training images might be required, at a minimum, to use this technique effectively. Internet searches suggested a minimum of around 1000 images per class, but he discovered that it was necessary to increase this number to 2500 to get better results. “My first approach,” he said, “was to develop a cross-platform app using technology such as React Native, which lets developers build apps for both iOS and Android using a common code base; but while dealing with deep learning models, this approach was limited. Instead, I had to develop custom iOS and Android apps separately to make the project work offline. It would be really nice to come up with a solution that uses a cross-mobile platform—such as React Native—and make it able to embed models for both iOS and Android offline use. That would definitely save lot of development time.”

For getting more deeply involved in an AI project, Kshitiz recommends visiting the Intel AI Developer Program. “The number of curated resources there is just mind blowing,” he said. “After that, free online courses are definitely something to consider, such as [DeepLearning.ai](#) from Coursera, [AI Student Courses from Intel AI Developer Program](#), and [Fast.ai](#) courses. I also recommend exploring the Intel Developer Mesh Platform to get inspiration and collaborate with fellow developers worldwide.”

Key insights that Kshitiz gained from the project include:

- Never trust the training and validation accuracy score; always try out your test data following training.
- Start with a small architecture model and gradually tweak the model as you increase or decrease the data.
- Don’t be impatient (like he says he was); use standard tests and procedures to discover better hyperparameters.
- First train on a moderate dataset. Then try to make sense of the predictions and increase or decrease the dataset as necessary.

AI is Expanding the Boundaries of Vision Applications

Through the design and development of specialized chips, research, educational outreach, and industry partnerships, Intel is accelerating the progress of AI to solve difficult challenges in medicine, manufacturing, agriculture, scientific research, robotics, and other industry sectors. Intel works closely with policymakers, educational institutions, and enterprises of all kinds to uncover and advance solutions that address major challenges in the sciences.

“If you think about just the evolutionary arc of humanity, information has been a function of the time we’ve lived in. If we were solving problems like shelter and food supply, the information we needed was simply the information to survive. Now that we’ve solved the big problems, we’ve started thinking about creative ways of solving more interesting problems.”²

- Naveen Rao, Vice President and General Manager, Artificial Intelligence Products Group, Intel

The Intel® AI portfolio includes:



Framework Optimization: Achieve faster training of deep neural networks on a robust scalable infrastructure.



Intel® Xeon® Scalable processors: Tackle AI challenges with a compute architecture optimized for a broad range of AI workloads, including deep learning.



Intel® Movidius™ Myriad™ X Vision Processing Unit (VPU): Delivers advanced features for the most demanding computer vision workloads and deep neural network implementations.



Intel® Movidius™ Neural Compute Stick: Provides deep learning prototyping at the network edge with always-on vision processing making it ideal for use in smart security cameras, gesture controlled drones, industrial machine vision equipment, and more.



Intel® Field Programmable Gate Arrays (FPGAs): Create specialized, custom functionality for a wide variety of electronic equipment, including AI-based solutions and monitoring devices, medical equipment, aircraft navigation devices, system accelerators, and more.



Reinforcement Learning Coach: Provides an open source research framework for training and evaluating RL agents by harnessing the power of multicore CPU processing to achieve state-of-the-art results.

Intel® Distribution of OpenVINO™ toolkit: Make your vision a reality on Intel® platforms—from smart cameras and video surveillance to robotics, transportation, and more.

Intel® Distribution for Python*: Supercharge applications and speed up core computational packages with this performance-oriented distribution.

Intel® Data Analytics Acceleration Library (Intel® DAAL): Boost machine learning and data analytics performance with this easy-to-use library.

Intel® Math Kernel Library (Intel® MKL): Accelerate math processing routines, increase application performance, and reduce development time.

For more information, visit the [portfolio page](#).

RESOURCES

Intel® AI Developer Program <https://software.intel.com/en-us/ai>

Inside Artificial Intelligence – Next-level computing powered by Intel AI ai.intel.com/

Intel® AI DevCloud – Free cloud compute for Intel AI Developer Program members
<https://software.intel.com/en-us/ai/devcloud>

Intel® Software Innovator Program - Supports innovative, independent developers
software.intel.com/intel-software-innovators/overview

Intel® Optimization for TensorFlow* ai.intel.com/tensorflow/

TensorFlow 501 – University Resources from Intel AI Developer Program
<https://software.intel.com/en-us/ai/courses/tensorflow>

Tips to Improve Performance for Popular Deep Learning Frameworks on CPUs
software.intel.com/en-us/articles/tips-to-improve-performance-for-popular-deep-learning-frameworks-on-multi-core-cpus

Keras*: The Python Deep Learning Library keras.io/

Artificial Intelligence for Development – A not-for-profit organization promoting research and development in AI
ainepal.org/mission/

Intel® Artificial Intelligence News newsroom.intel.com/press-kits/artificial-intelligence/



¹ "The Rise of Artificial Intelligence for the Visually Impaired." The Vision of Children Foundation. April 2018.
<https://www.visionofchildren.org/voc-blog/2018/4/12/the-rise-of-artificial-intelligence-for-the-visually-impaired>

² Bhushan, Kul. "We're not even close building a true AI: Intel's Naveen Rao." LiveMint. August 2018.
<https://www.livemint.com/AI/79F1MOYQoWdWWRs4TBXbLM/Were-not-even-close-to-building-a-true-Artificial-Intelligence.html>

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications, and roadmaps.

The products and services described may contain defects or errors known as errata which may cause deviations from published specifications. Current characterized errata are available on request.

Copies of documents which have an order number and are referenced in this document may be obtained by calling 1-800-548-4725 or by visiting www.intel.com/design/literature.htm.

Intel, the Intel logo, OpenVINO, Movidius, and Xeon are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.

© 2018 Intel Corporation Printed in USA 1218/KP/MESH/PDF 338130-001US