ASEE AMERICAN SOCIETY FOR ENGINEERING EDUCATION

Paper ID #35976

Using Neural Networks to Distinguish Children's Age with Visual Features of Sketches

Mr. Aniket Patel, Texas A&M University

Aniket Patel is a junior in Computer Science at Texas A&M University. He is working as an undergraduate researcher pursuing how children's drawing ability links to other developmental features associated with learning and how machine learning can be applied to this space. He previously worked as a researcher studying material science and analyzed material diffraction patterns.

Mr. Seth Polsley, Texas A&M University

Seth Polsley is a PhD student at Texas A&M University in the Sketch Recognition Lab under Director Tracy Hammond. His research interests may be broadly classified as "intelligent systems," with an emphasis on studying and building interactions that merge the capabilities of computers with the intuitive behaviors of humans. He holds a Masters and Bachelors in Computer Engineering from Texas A&M and University of Kansas, respectively, and has previously worked at Lexmark International and MIT Lincoln Lab.

Dr. Tracy Anne Hammond, Texas A&M University

Dr. Hammond is Director of the Texas A&M University Institute for Engineering Education & Innovation and also the chair of the Engineering Education Faculty. She is also Director of the Sketch Recognition Lab and Professor in the Department of Computer Science & Engineering. She is a member of the Center for Population and Aging, the Center for Remote Health Technologies & Systems as well as the Institute for Data Science. Hammond is a PI for over 13 million in funded research, from NSF, DARPA, Google, Microsoft, and others. Hammond holds a Ph.D. in Computer Science and FTO (Finance Technology Option) from the Massachusetts Institute of Technology, and four degrees from Columbia University: an M.S in Anthropology, an M.S. in Computer Science, a B.A. in Mathematics, and a B.S. in Applied Mathematics and Physics. Hammond advised 17 UG theses, 29 MS theses, and 10 Ph.D. dissertations. Hammond is the 2020 recipient of the TEES Faculty Fellows Award and the 2011 recipient of the Charles H. Barclay, Jr. '45 Faculty Fellow Award. Hammond has been featured on the Discovery Channel and other news sources. Hammond is dedicated to diversity and equity, which is reflected in her publications, research, teaching, service, and mentoring. More at http://srl.tamu.edu and http://ieei.tamu.edu.

Session XXXX

Using Neural Networks to Distinguish Children's Age with Visual Features of Sketches

Aniket Patel, Seth Polsley, Tracy Hammond
Department of Computer Science and Engineering
Texas A&M University

Abstract

Children's fine motor control is linked to other critical skills during early childhood development, such as school-readiness and reading comprehension. However, many assessments require expert evaluation or can be prone to bias. Through machine learning, we seek to provide parents and teachers with easier access to fine motor skill assessment. In this work, we develop a vision-based approach using neural networks to predict children's age groups by considering only their drawings. While it does not support detailed fine motor analysis yet, this effort is part of a larger project to apply sketch recognition to assess children's fine motor ability, and we focus on vision-based assessment to distinguish approximate ages in order to compare with human expert evaluation on the same images. Our neural network achieved generalization and validation accuracies of 75.0%, with a set of curve drawings, and a generalization and validation accuracy of 72.3% with the corner dataset. By comparison, we also asked two human evaluators to label age groups based on a portion of the dataset, and they achieved 77.5% and 66.7% for the curve and corner drawings, respectively. These findings provide support for future work integrating more sketch recognition features, such as pressure and timing information, in order to make better assessments that are fast, easy, and accessible for parents and children.

Introduction

Children's fine motor control is linked to other critical skills during early childhood development, including school-readiness, reading comprehension, and math performance[1]. Detecting fine motor delay is important to ensure healthy development and avoid delays in other areas[2].

Experts use a variety of methods to assess fine motor control. Many of these methods require in-person assessments to be done by a pediatrician or other child development expert. However, there are some popular techniques that enable parents to make assessments at home, such as the Ages and Stages Questionnaire[3]. Indirect assessments such as this come at a greater availability and lower cost, but are also not as accurate as direct assessments by an expert. Several of these assessment techniques make use of drawing, among other activities, as it is linked to fine motor ability and later achievement[4].

While there are some options for determining fine motor control development, most assessments that use sketching rely on shape correctness. However, this requires a human expert to evaluate the drawings and uses only visual features, introducing potential for bias or errors[5,6]. Furthermore, due to lack of awareness, costs associated with assessments, and accessibility issues, parents may not know that their child is developmentally delayed until additional therapy is needed[7]. We take a computer vision approach in hopes to combat the shortcomings of both indirect and direct assessment techniques, and provide an accurate, unbiased, and easily available assessment technique to assess fine motor control.

Computer vision is a field of artificial intelligence focused on training computers to analyze visual inputs such as digital pictures or videos and take actions based on that information, namely classification. It tries to replicate the human sight by recognizing patterns and making decisions based on that. It is being used in numerous fields including education.

Methodology

Our data set consisted of a collection of digital drawings from 60 children aged 3-8 years old. This included 6 total drawings, 293 drawings from children aged 3–4 and 258 drawings from children aged 5–8. The children, while being accompanied by their parents, were given a prompt to draw a specific shape or letter on a tablet using a digital stylus. The collection of drawings was divided into two subsets: a corner subset, which involved drawings such as the letter ‘A’ or a triangle, as shown in Figure 2; and a curve subset which involved drawings such as the letter ‘C’ or a circle, as shown in Figure 3. As explained in a previous study[7], while children are able to recognise certain shapes, such as triangles, squares, and rectangles, in their primary forms, the capability of recognising shapes in an altered form is first seen from around the age of 5. Due to this, we selected the age groups to be split into children under 5 and those 5 or older.

Figure 1: Examples of the resized, binary digital drawings. Image on the left is an example of a drawing from the corner subset, while the image on the right is an example of a

drawing from the curve subset.

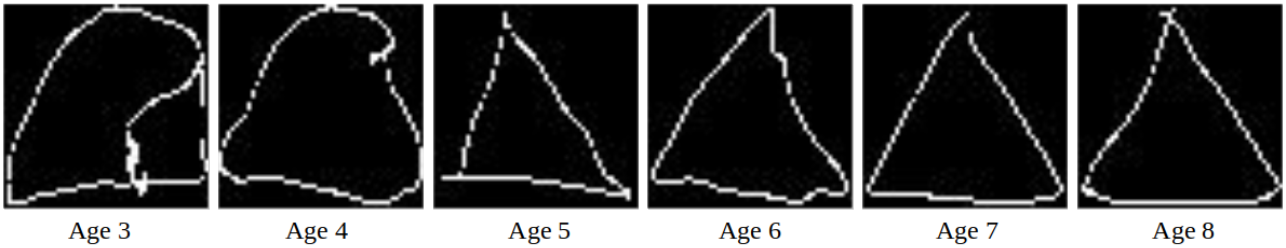

Figure 2: Examples of drawings from children given the prompt to draw a triangle (corner shape).

Figure 3: Examples of drawings from children given the prompt to draw a circle (curve shape).

Each digital drawing was saved as a list of strokes in a .xml file. Every stroke was a collection of points with the x-position, y-position, and time saved. Using the x and y position data from the file and the Python package matplotlib[8], each file was converted back into a digital drawing.

The vision processing method we are using is a feed-forward, back-propagation neural network implemented in Python. A feed-forward neural network is simply a network that passes information in one direction without any loops. In the context of this project, feed-forward means that the input data is fed into the input layer which passes it to the hidden layer and then finally to the output layer. Back-propagation refers to the algorithm used to adjust the weights of the neural network in order to increase the accuracy of predictions made. The algorithm works by trying to minimize the loss function with respect to the weights, iterating backwards from the output layer and adjusting the weights to an appropriate degree to attain the desired output over several iterations.

The network considers digital drawings as a collection of pixels to make classification decisions based only on the appearance of a drawing. To achieve this, the image is down-scaled to 50x50 and converted to a binary scale from greyscale. This means that every pixel in the image is characterized by a single bit; where a 1 corresponds to a white pixel and a 0 corresponds to a black pixel. Next, the 50x50 drawings are flattened into a 1 dimensional array which is then written to an input file in a random ordering. Randomizing the order the drawings are passed into the neural network can prevent the network from learning the order of training data and reduce bias. This means predictions will be based on the input data rather than the order of the input data, resulting in a better generalization accuracy.

For the classification output, the network has two decisions: 0 for children ages 3–4 and 1 for children ages 5–8, which are written to the output file in the matching random order. This generates an input file where each line contains the flattened array of a single digital drawing and the output file contains the correct classification of that digital image in that same corresponding line.

The neural network we are using has an input layer with 2500 nodes, a hidden layer with 200 nodes and an output layer with 2 nodes. The network reads in the data from the input file line by line as each line has the flattened array containing the digital drawing. At the end of the training segment, the performance is evaluated using a portion of the data set isolated for generalization purposes, and after all epochs are complete, the validation test set results are evaluated using standard performance metrics. 60% of the input data was used to train the network, while the other 40% of the data goes towards generalizing and validation equally.

Results

The network outputs a 0 if it classifies the digital drawing as a drawing from a child aged 3–4 and outputs a 1 if the network classifies it as a drawing from a child aged 5-8. Our neural network achieved generalization and validation accuracies of 75.0%, with a set of curve drawings, and a generalization and validation accuracy of 72.3% with the corner dataset. A previous study that used human evaluators to classify the digital drawings achieved 77.5% in the curve dataset and 66.7% in the corner dataset[7].

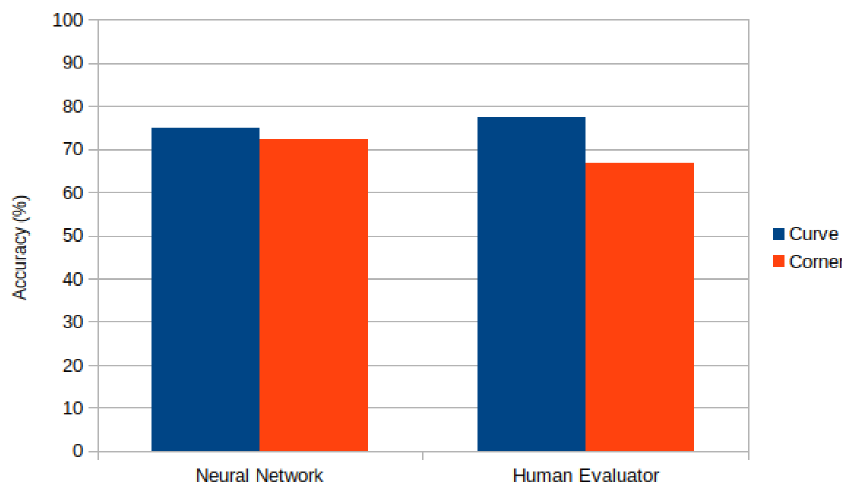

Figure 4: Graph of the comparative accuracy between the neural network and human evaluators.

As in the previous study that considered only sketch-based features[7], the human evaluators outperformed our vision-based neural network in the curve subset, but fell behind in the corner subset. Perhaps with a deeper neural network and more training data or more features, the neural network can significantly improve to outperform the human evaluators in both the categories.

Conclusion

In this paper, we illustrated a computer vision approach to classify children according to fine motor function with relation to age using digital drawings. We implemented a neural network that classified a digital drawing as either from a child aged 3–4 or a child aged 5 and over. We compared the performance of our neural network on the curve and corner subsets to that of human evaluators and were able to achieve similar or slightly better results. Our next steps include building on existing sketch recognition research and incorporating more features such as tilt and pressure measurements in hopes of building a reliably accurate application to access children's fine motor skills.

References

1. Grissmer, D., Grimm, K.J., Aiyer, S.M., Murrah, W.M., Steele, J.S.: Fine motor skills and early comprehension of the world: two new school readiness indicators. Develop-mental psychology 46(5), 1008 (2010).
2. Anthony, L., Brown, Q., Nias, J., Tate, B., Mohan, S.: Interaction and recognition challenges in interpreting children's touch and gesture input on mobile devices. In:Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces, pp. 225–234 (2012).
3. Bricker, D., Squires, J., Mounts, L., Potter, L., Nickel, R., Twombly, E., & Farrell, J. Ages and stages questionnaire. Paul H. Brookes: Baltimore. (1999).
4. Suggate, S., Pufke, E., Stoeger, H.: Do fine motor skills contribute to early reading development? Journal of Research in Reading 41(1), 1–19 (2018)
5. Kim, H., Taele, P., Valentine, S., McTigue, E., Hammond, T.: Kimchi: a sketch-based developmental skill classifier to enhance pen-driven educational interfaces for children.In: Proceeding SBIM´13 Proceedings of the International Symposium on Sketch-BasedInterfaces and Modeling, pp. 33–42 (2013).
6. Lotz, L., Loxton, H., Naidoo, A.: Visual-motor integration functioning in a south african middle childhood sample. Journal of Child & Adolescent Mental Health pp. 63–67 (2005).
7. Polsley, S., Powell, L., Kim, H., Thomas, X., Liew, J., and Hammond T.: Detecting Children's Fine Motor Skill Development using Machine Learning. International Journal of Artificial Intelligence in Education 1-34 (2021).
8. Barrett, Paul, John Hunter, J. Todd Miller, J-C. Hsu, and Perry Greenfield. "matplotlib--A Portable Python Plotting Package." In Astronomical data analysis software and systems XIV, vol. 347, p. 91. 2005.

ANIKET PATEL
Aniket Patel is a junior in Computer Science at Texas A&M University. He is an undergraduate researcher pursuing how children's drawing ability links to other developmental features tied to learning and how machine learning can be applied to this space. He previously worked as a researcher studying material science, analyzing material diffraction patterns.

SETH POLSLEY
Seth Polsley is a PhD student at Texas A&M University in the Sketch Recognition Lab. His research interests may be broadly classified as "intelligent systems," with an emphasis on interactions that merge the capabilities of computers with intuitive behaviors of humans. He holds a Masters and Bachelors in Computer Engineering from Texas A&M and University of Kansas, respectively, and has previously worked at Lexmark

and MIT Lincoln Lab.

TRACY HAMMOND
Dr. Tracy Hammond is Director of the Institute for Engineering Education and Innovation, Director of the Sketch Recognition Lab, and Professor in the Department of Computer Science and Engineering at Texas A&M. She is an international leader in sketch recognition and human-computer interaction research. Her research has been funded by NSF, DARPA, Google, Microsoft, and many others. She holds a Ph.D. in Computer Science and Finance Technology Option from M.I.T., and an M.S. in Anthropology, M.S. in Computer Science, B.A. in Mathematics, and B.S. in Applied Mathematics from Columbia University.