THE WITS INTELLIGENT TEACHING SYSTEM (WITS)
A SMART LECTURE THEATRE TO ASSESS AUDIENCE ENGAGEMENT

SCHOOL OF COMPUTER SCIENCE AND APPLIED MATHEMATICS
UNIVERSITY OF THE WITWATERSRAND

RICHARD KLEIN
0707074G

SUPERVISED BY
PROF TURGAY CELIK

MARCH 6, 2018

A Thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg,
in fulfilment of the requirements for the degree of Doctor of Philosophy

Ethics Clearance Number: H14/03/06

Abstract

The utility of lectures is directly related to the engagement of the students therein. To ensure the value of lectures, one needs to be certain that they are engaging to students. In small classes experienced lecturers develop an intuition of how engaged the class is as a whole and can then react appropriately to remedy the situation through various strategies such as breaks or changes in style, pace and content. As both the number of students and size of the venue grow, this type of contingent teaching becomes increasingly difficult and less precise. Furthermore, relying on intuition alone gives no way to recall and analyse previous classes or to objectively investigate trends over time. To address these problems this thesis presents the WITS INTELLIGENT TEACHING SYSTEM (WITS) to highlight disengaged students during class.

A web-based, mobile application called Engage was developed to try elicit anonymous engagement information directly from students. The majority of students were unwilling or unable to self-report their engagement levels during class. This stems from a number of cultural and practical issues related to social display rules, unreliable internet connections, data costs, and distractions. This result highlights the need for a non-intrusive system that does not require the active participation of students. A non-intrusive, computer vision and machine learning based approach is therefore proposed.

To support the development thereof, a labelled video dataset of students was built by recording a number of first year lectures. Students were labelled across a number of affects – including boredom, frustration, confusion, and fatigue – but poor inter-rater reliability meant that these labels could not be used as ground truth. Based on manual coding methods identified in the literature, a number of actions, gestures, and postures were identified as proxies of behavioural engagement. These proxies are then used in an observational checklist to mark students as engaged or not.

A Support Vector Machine (SVM) was trained on Histograms of Oriented Gradients (HOG) to classify the students based on the identified behaviours. The results suggest a high temporal correlation of a single subject's video frames. This leads to extremely high accuracies on seen subjects. However, this approach generalised poorly to unseen subjects and more careful feature engineering is required. The use of Convolutional Neural Networks (CNNs) improved the classification accuracy substantially, both over a single subject and when generalising to unseen subjects. While more computationally expensive than the SVM, the CNN approach lends itself to parallelism using Graphics Processing Units (GPUs). With GPU hardware acceleration, the system is able to run in near real-time and with further optimisations a real-time classifier is feasible.

The classifier provides engagement values, which can be displayed to the lecturer live during class. This information is displayed as an Interest Map which highlights spatial areas of disengagement. The lecturer can then make informed decisions about how to progress with the class, what teaching styles to employ, and on which students to focus. An Interest Map was presented to lecturers and professors at the University of the Witwatersrand yielding 131 responses. The vast majority of respondents indicated that they would like to receive live engagement feedback during class, that they found the Interest Map an intuitive visualisation tool, and that they would be interested in using such technology.

Contributions of this thesis include the development of a labelled video dataset; the development of a web based system to allow students to self-report engagement; the development of cross-platform, open-source software for spatial, action and affect labelling; the application of Histogram of Oriented Gradient based Support Vector Machines, and Deep Convolutional Neural Networks to classify this data; the development of an Interest Map to intuitively display engagement information to presenters; and finally an analysis of acceptance of such a system by educators.

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

March 6, 2018

Dedication

Acknowledgements

Ack.

Contents

List of Figures

List of Tables

Nomenclature

Mathematical Notation

$\mathbf{A}, \mathbf{B}, \mathbf{W}$ An upper case, boldface letter is a matrix.

X, Y, Z An upper case, light (non-boldface) letter is a set.

Images

$\mathbb{D}$ A dataset.

$\mathbf{I}$ An image or image sequence in dataset $\mathbb{D}$.

$\mathbf{I}(x, y)$ The pixel value at position (x, y) of some image $\mathbf{I}$.

Chapter 1

Introduction

1.1 Introduction

S TUDENT engagement is a pivotal concept when considering the utility of lectures. In an age
where even tertiary level content is available online, it becomes important for instructors
to differentiate themselves from online content sources. This is particularly true when teaching
large classes where interaction with the presenter is already limited.

```cpp
int main(){
    cout << "Hello World" << endl;
}
```

References

[Adair 1984] John G Adair. The Hawthorne effect: A reconsideration of the methodological artifact. *Journal of applied psychology*, 69(2):334, 1984.

[Adiv 1985] Gilad Adiv. Determining three-dimensional motion and structure from optical flow generated by several moving objects. *IEEE transactions on pattern analysis and machine intelligence*, (4):384–401, 1985.

[Afzal and Robinson 2009] Shazia Afzal and Peter Robinson. Natural affect data—collection & annotation in a learning context. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–7. IEEE, 2009.

[Agrawal *et al.* 1993] Rakesh Agrawal, Christos Faloutsos, and Arun Swami. Efficient similarity search in sequence databases. In *International Conference on Foundations of Data Organization and Algorithms*, pages 69–84. Springer, 1993.

[Ahonen *et al.* 2006] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, Dec 2006.

[Alyuz *et al.* 2016] Nese Alyuz, Eda Okur, Ece Oktay, Utku Genc, Sinem Aslan, Sinem Emine Mete, David Stanhill, Bert Arnrich, and Asli Arslan Esme. Towards an emotional engagement model: Can affective states of a learner be automatically detected in a 1: 1 learning scenario. In *Proceedings of the 6th Workshop on Personalization Approaches in Learning Environments (PALE 2016)*. 24th conference on User Modeling, Adaptation, and Personalization (UMAP 2016), CEUR workshop proceedings, this volume, 2016.

[Anderson *et al.* 2004] Amy R Anderson, Sandra L Christenson, Mary F Sinclair, and Camilla A Lehr. Check & Connect: The importance of relationships for promoting engagement with school. *Journal of School Psychology*, 42(2):95–113, 2004.

[Andrade 2010] Jackie Andrade. What Does Doodling do? *Applied Cognitive Psychology*, 24:100–106, February 2010.

[Arroyo *et al.* 2004] Ivon Arroyo, Carole Beal, Tom Murray, Rena Walles, and Beverly P Woolf. Web-based intelligent multimedia tutoring for high stakes achievement tests. In *Intelligent Tutoring Systems*, pages 468–477. Springer, 2004.

[Arroyo *et al.* 2007] Ivon Arroyo, Kimberly Ferguson, Jeffrey Johns, Toby Dragon, Hasmik Meheranian, Don Fisher, Andrew Barto, Sridhar Mahadevan, and Beverly Park Woolf. Repairing disengagement with non-invasive interventions. In *AIED*, volume 2007, pages 195–202, 2007.

[Arroyo *et al.* 2009] Ivon Arroyo, David G Cooper, Winslow Burleson, Beverly Park Woolf, Kasia Muldner, and Robert Christopherson. Emotion Sensors Go To School. In *AIED*, volume 200, pages 17–24, 2009.

[Asthana *et al.* 2009] Akshay Asthana, Jason Saragih, Michael Wagner, and Roland Goecke. Evaluating aam fitting methods for facial expression recognition. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–8. IEEE, 2009.

[Astin 1984] Alexander W Astin. Student involvement: A developmental theory for higher education. *Journal of college student personnel*, 25(4):297–308, 1984.

[Averill 1980] J.R Averill. A constructivist view of emotion. In R. Plutchik and H. Kellerman, editors, *Emotion: Theory, research and experience*, pages 305–339. Academic Press, 1980.

[Ba and Odobez 2009] Sileye O Ba and J-M Odobez. Recognizing visual focus of attention from head pose in natural meetings. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 39(1):16–33, 2009.

[Ba and Odobez 2011] Sileye O Ba and J Odobez. Multiperson visual focus of attention from head pose and meeting contextual cues. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(1):101–116, 2011.

[Ba 2007] S. Ba. *Joint Head Tracking and Pose Estimation for Visual Focus of Attention Recognition,*. PhD thesis, École Polytechnique Fédérale de Lausanne (EPFL), February 2007.

[Barrett 2006] Lisa Feldman Barrett. Are emotions natural kinds? *Perspectives on psychological science*, 1(1):28–58, 2006.

[Benzaid and Dewan 2010] Sami Benzaid and Prasun Dewan. Semantic Awareness Through Computer Vision. In *Proceedings of the 2nd ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, EICS '10, pages 205–210, New York, NY, USA, 2010. ACM.

[Bianchi-Berthouze and Lisetti 2002] Nadia Bianchi-Berthouze and Christine L Lisetti. Modeling multimodal expression of user's affective subjective experience. *User Modeling and User-Adapted Interaction*, 12(1):49–84, 2002.

[Blaszczynski *et al.* 1990] Alex Blaszczynski, Neil McConaghy, and Anna Frankova. Boredom Proneness in Pathological Gambling. *Psychological Reports*, 67(1):35–42, 1990.

[Bosch *et al.* 2015a] Nigel Bosch, Huili Chen, Sidney D'Mello, Ryan Baker, and Valerie Shute. Accuracy vs. availability heuristic in multimodal affect detection in the wild. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, pages 267–274. ACM, 2015.

[Bosch *et al.* 2015b] Nigel Bosch, Sidney D'Mello, Ryan Baker, Jaclyn Ocumpaugh, Valerie Shute, Matthew Ventura, Lubin Wang, and Weinan Zhao. Automatic Detection of Learning-Centered Affective States in the Wild. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*, IUI '15, pages 379–388, New York, NY, USA, 2015. ACM.

[Boser *et al.* 1992] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152. ACM, 1992.

[Bouguet 2001] Jean-Yves Bouguet. Pyramidal implementation of the affine Lucas-Kanade feature tracker description of the algorithm. *Intel Corporation*, 5(1-10):4, 2001.

[Boyle and Nicol 2003] James T Boyle and David J Nicol. Using classroom communication systems to support interaction and discussion in large class settings. *Research in Learning Technology*, 11(3), 2003.

[Bradski 2000] Gary Bradski. The OpenCV library. *Doctor Dobbs Journal*, 25(11):120–126, 2000.

[Brett 2011] Paul Brett. Students' experiences and engagement with SMS for learning in higher education. *Innovations in Education and Teaching International*, 48(2):137–147, 2011.

[Brick *et al.* 2009] Timothy R Brick, Michael D Hunter, and Jeffrey F Cohn. Get The FACS Fast: Automated FACS face analysis benefits from the addition of velocity. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–7. IEEE, 2009.

[Bull 1987] Peter E Bull. *Posture and gesture.* Pergamon press, 1987.

[Calvo and D'Mello 2010] Rafael A Calvo and Sidney D'Mello. Affect detection: An interdisciplinary review of models, methods, and their applications. *Affective Computing, IEEE Transactions on*, 1(1):18–37, 2010.

[Calvo and Nummenmaa 2007] Manuel G Calvo and Lauri Nummenmaa. Processing of unattended emotional visual scenes. *Journal of Experimental Psychology: General*, 136(3):347, 2007.

[Card *et al.* 1974] WI Card, Mary Nicholson, GP Crean, Geoffrey Watkinson, CR Evans, Jackie Wilson, and Daphne Russell. A comparison of doctor and computer interrogation of patients. *International journal of bio-medical computing*, 5(3):175–187, 1974.

[Castellano *et al.* 2008a] Ginevra Castellano, Loic Kessous, and George Caridakis. Emotion recognition through multiple modalities: face, body gesture, speech. In *Affect and emotion in human-computer interaction*, pages 92–103. Springer, 2008.

[Castellano *et al.* 2008b] Ginevra Castellano, Marcello Mortillaro, Antonio Camurri, Gualtiero Volpe, and Klaus Scherer. Automated analysis of body movement in emotionally expressive piano performances. *Music Perception: An Interdisciplinary Journal*, 26(2):103–119, 2008.

[Chan and Fu 1999] Kin-Pong Chan and Ada Wai-Chee Fu. Efficient time series matching by wavelets. In *Data Engineering, 1999. Proceedings., 15th International Conference on*, pages 126–133. IEEE, 1999.

[Chang and Lin 2011] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[Chang *et al.* 2002] Jyh-Yeong Chang, Wen-Feng Hu, Mu-Huo Cheng, and Bo-Sen Chang. Digital image translational and rotational motion stabilization using optical flow technique. *IEEE Transactions on Consumer Electronics*, 48(1):108–115, 2002.

[Chapelle 2007] Olivier Chapelle. Training a support vector machine in the primal. *Neural computation*, 19(5):1155–1178, 2007.

[Chen and Chung 2008] Chih-Ming Chen and Ching-Ju Chung. Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education*, 51(2):624–645, 2008.

[Chetlur *et al.* 2014] Sharan Chetlur, Cliff Woolley, Philippe Vandermersch, Jonathan Cohen, John Tran, Bryan Catanzaro, and Evan Shelhamer. cudnn: Efficient primitives for deep learning. *CoRR*, abs/1410.0759, 2014.

[Cohen 1960] Jacob Cohen. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, (20):37–46, 1960.

[Cohn and Schmidt 2004] Jeffrey F Cohn and Karen L Schmidt. The timing of facial motion in posed and spontaneous smiles. *International Journal of Wavelets, Multiresolution and Information Processing*, 2(2):121–132, 2004.

[Conati and Maclare 2004] Cristina Conati and Heather Maclare. Evaluating a probabilistic model of student affect. In *Intelligent tutoring systems*, pages 55–66. Springer, 2004.

[Cooper 1960] L. Cooper. *Aristotle, The Rhetoric of Aristotle*. Appleton-Century-Crofts, 1960. An expanded translation with supplementary examples for students of composition and public speaking.

[Coulson 2004] Mark Coulson. Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of nonverbal behavior*, 28(2):117–139, 2004.

[Craig *et al.* 2008] Scotty D. Craig, Sidney D'Mello, Amy Witherspoon, and Art Graesser. Emote aloud during learning with AutoTutor: Applying the Facial Action Coding System to cognitive–affective states during learning. *Cognition & Emotion*, 22(5):777–788, 2008.

[Cytowic 1996] Richard E Cytowic. *The neurological side of neuropsychology*. MIT Press, 1996.

[Dalal and Triggs 2005] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.

[Dalgleish *et al.* 2009] Tim Dalgleish, Barnaby D Dunn, and Dean Mobbs. Affective neuroscience: Past, present, and future. *Emotion Review*, 1(4):355–368, 2009.

[Dalgleish 2004] Tim Dalgleish. The emotional brain. *Nature Reviews Neuroscience*, 5(7):583–589, 2004.

[Damasio 1994] Antonio R Damasio. *Descartes' error: Emotion, reason, and the human brain*, 1994.

[Damasio 2000] Antonio Damasio. *Looking for Spinoza. Joy, Sorrow, and the Feeling Brain*, 2000.

[Darwin 1859] Charles Darwin. On the origins of species by means of natural selection. *London: Murray*, 1859.

[Darwin 1872] Charles Darwin. *The Expression of the Emotions in Man and Animals*. John Murray, 1872.

[Davidson *et al.* 2003] Richard J Davidson, Klaus R Scherer, and Hill Goldsmith. *Handbook of affective sciences*. Oxford University Press, 2003.

[De Meijer 1989] Marco De Meijer. The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal behavior*, 13(4):247–268, 1989.

[Dearnley *et al.* 2008] Chris Dearnley, Jackie Haigh, and John Fairhall. Using mobile technologies for assessment and learning in practice settings: a case study. *Nurse education in practice*, 8(3):197–204, 2008.

[Dennerlein *et al.* 2003] Jack Dennerlein, Theodore Becker, Peter Johnson, Carson Reynolds, and Rosalind W Picard. Frustrating computer users increases exposure to physical factors. In *Proceedings of the International Ergonomics Association, Seoul, Korea*, 2003.

[Devi and Bajaj 2008] M.S. Devi and P.R. Bajaj. Driver Fatigue Detection Based on Eye Tracking. In *Emerging Trends in Engineering and Technology, 2008. ICETET '08. First International Conference on*, pages 649–652, July 2008.

[D'mello and Graesser 2007] Sidney D'mello and Arthur Graesser. Mind and Body: Dialogue and posture for affect detection in learning environments. *Frontiers in Artificial Intelligence and Applications*, 158:161, 2007.

[D'Mello and Graesser 2009] Sidney D'Mello and Art Graesser. Automatic Detection of Learner's Affect From Gross Body Language. *Applied Artificial Intelligence*, 23(2):123–150, February 2009.

[D'Mello and Graesser 2010] Sidney K D'Mello and Arthur Graesser. Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-Adapted Interaction*, 20(2):147–187, 2010.

[D'Mello *et al.* 2006] Sidney K D'Mello, Scotty D Craig, Jeremiah Sullins, and Arthur C Graesser. Predicting affective states expressed through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue. *International Journal of Artificial Intelligence in Education*, 16(1):3–28, 2006.

[D'Mello *et al.* 2007] Sidney D'Mello, Rosalind Picard, and Arthur Graesser. Towards an affect-sensitive autotutor. *IEEE Intelligent Systems*, 22(4):53–61, 2007.

[D'Mello *et al.* 2010] Sidney D'Mello, Blair Lehman, Jeremiah Sullins, Rosaire Daigle, Rebekah Combs, Kimberly Vogt, Lydia Perkins, and Art Graesser. A time for emoting: When affect-sensitivity is and isn't effective at promoting deep learning. In *Intelligent tutoring systems*, pages 245–254. Springer, 2010.

[Donato *et al.* 1999] Gianluca Donato, Marian Stewart Bartlett, Joseph C. Hager, Paul Ekman, and Terrence J. Sejnowski. Classifying facial actions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(10):974–989, 1999.

[Dong and Wu 2005] Wenhui Dong and Xiaojuan Wu. Fatigue detection based on the distance of eyelid. In *VLSI Design and Video Technology, 2005. Proceedings of 2005 IEEE International Workshop on*, pages 365–368, May 2005.

[Draper and Brown 2004] Stephen W Draper and Margaret I Brown. Increasing interactivity in lectures using an electronic voting system. *Journal of computer assisted learning*, 20(2):81–94, 2004.

[DMello 2013] Sidney DMello. A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology*, 105(4):1082, 2013.

[Ekman and Friesen 1969] Paul Ekman and Wallace V Friesen. *Nonverbal leakage and clues to deception*. Technical report, DTIC Document, 1969.

[Ekman and Friesen 1978] Paul Ekman and Wallace V Friesen. Facial action coding system: A technique for the Measurement of Facial Movement: Investigators Guide 2 Parts. *Consulting Psychologists Press*, 1978.

[Ekman and Friesen 2003] Paul Ekman and Wallace V Friesen. *Unmasking the face: A guide to recognizing emotions from facial clues*. Ishk, 2003.

[Ekman *et al.* 1980] Paul Ekman, Wallace V Freisen, and Sonia Ancoli. Facial signs of emotional experience. *Journal of personality and social psychology*, 39(6):1125, 1980.

[Ekman *et al.* 1990] Paul Ekman, Richard J Davidson, and Wallace V Friesen. The Duchenne smile: Emotional expression and brain physiology: II. *Journal of personality and social psychology*, 58(2):342, 1990.

[Ekman 1971] Paul Ekman. Universals and cultural differences in facial expressions of emotion. In *Nebraska symposium on motivation*. University of Nebraska Press, 1971.

[Ekman 1992] Paul Ekman. An argument for basic emotions. *Cognition & Emotion*, 6(3-4):169–200, 1992.

[El Kaliouby and Robinson 2005a] Rana El Kaliouby and Peter Robinson. Generalization of a vision-based computational model of mind-reading. In *Affective computing and intelligent interaction*, pages 582–589. Springer, 2005.

[El Kaliouby and Robinson 2005b] Rana El Kaliouby and Peter Robinson. Real-time inference of complex mental states from facial expressions and head gestures. In *Real-time vision for human-computer interaction*, pages 181–200. Springer, 2005.

[Faloutsos *et al.* 1994] Christos Faloutsos, Mudumbai Ranganathan, and Yannis Manolopoulos. *Fast subsequence matching in time-series databases*, volume 23. ACM, 1994.

[Faure and Orthober 2011] Caroline Faure and Corrie Orthober. Using text-messaging in the secondary classroom. *American Secondary Education*, 39(2):55, 2011.

[Fecci *et al.* 1971] R Fecci, R Bartelemy, J Bourgoin, A Mathia, H Eberle, A Moutel, and G Jullien. Effects of infrasound on the organism. *La Medicina del Lavoro*, 62:130–150, 1971.

[Freund and Schapire 1995] Yoav Freund and Robert E Schapire. A desicion-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory*, pages 23–37. Springer, 1995.

[Friedman *et al.* 2001] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics Springer, Berlin, 2001.

[Frijda 1987] Nico H Frijda. Emotion, cognitive structure, and action tendency. *Cognition and emotion*, 1(2):115–143, 1987.

[Gavrila 1999] Dariu M Gavrila. The visual analysis of human movement: A survey. *Computer vision and image understanding*, 73(1):82–98, 1999.

[Gendron *et al.* 2012] Maria Gendron, Kristen A Lindquist, Lawrence Barsalou, and Lisa Feldman Barrett. Emotion words shape emotion percepts. *Emotion*, 12(2):314, 2012.

[Glenberg *et al.* 2005] A.M. Glenberg, D Havas, R Becker, and M Rinck. Grounding language in bodily states. In D Pecher and R.A. Zwaan, editors, *Grounding cognition: The role of perception and action in memory, language, and thinking*, pages 115–128. Cambridge University Press, Cambridge, 2005.

[Go *et al.* 2009] Alec Go, Lei Huang, and Richa Bhayani. Twitter sentiment analysis. *Entropy*, 17, 2009.

[Gogia *et al.* 2016] Yash Gogia, Eejya Singh, Shreyash Mohatta, and V Sreejith. Multi-modal affect detection for learning applications. In *Region 10 Conference (TENCON), 2016 IEEE*, pages 3743–3747. IEEE, 2016.

[Goodfellow *et al.* 2016] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[Graesser *et al.* 2006] AC Graesser, Bethany McDaniel, Patrick Chipman, Amy Witherspoon, Sidney D'Mello, and Barry Gholson. Detection of emotions during learning with AutoTutor. In *Proceedings of the 28th Annual Meetings of the Cognitive Science Society*, pages 285–290. Citeseer, 2006.

[Graesser *et al.* 2007] Arthur Graesser, Patrick Chipman, Brandon King, Bethany McDaniel, and Sidney D'Mello. Emotions and learning with auto tutor. *Frontiers in Artificial Intelligence and Applications*, 158:569, 2007.

[Graham and Weiner 1996] Sandra Graham and Bernard Weiner. Theories and principles of motivation. *Handbook of educational psychology*, 4:63–84, 1996.

[Haro *et al.* 2000] Antonio Haro, Myron Flickner, and Irfan Essa. Detecting and tracking eyes by using their physiological properties, dynamics, and appearance. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, pages 163–168. IEEE, 2000.

[Healey and Picard 2000] Jennifer Healey and Rosalind Picard. Smartcar: detecting driver stress. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 4, pages 218–221. IEEE, 2000.

[Heisele *et al.* 2001] B. Heisele, P. Ho, and T. Poggio. Face recognition with support vector machines: global versus component-based approach. In *Proceedings of the Eighth IEEE International Conference on Computer Vision (ICCV2001)*, volume 2, pages 688–694, 2001.

[Heudorf *et al.* 2009] U. Heudorf, V. Neitzert, and J. Spark. Particulate matter and carbon dioxide in classrooms – The impact of cleaning and ventilation. *International Journal of Hygiene and Environmental Health*, 212(1):45–55, 2009.

[Hinton *et al.* 2012] Geoffrey E Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan R Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*, 2012.

[Holley and Oliver 2010] Debbie Holley and Martin Oliver. Student engagement and blended learning: Portraits of risk. *Computers & Education*, 54(3):693–700, 2010.

[Hoque *et al.* 2009] Mohammed E Hoque, Rana El Kaliouby, and Rosalind W Picard. When human coders (and machines) disagree on the meaning of facial affect in spontaneous videos. In *Intelligent Virtual Agents*, pages 337–343. Springer, 2009.

[Horn and Schunck 1981] Berthold KP Horn and Brian G Schunck. Determining optical flow. *Artificial intelligence*, 17(1-3):185–203, 1981.

[Horng *et al.* 2004] Wen-Bing Horng, Chih-Yuan Chen, Yi Chang, and Chun-Hai Fan. Driver fatigue detection based on eye tracking and dynamk, template matching. In *Networking, Sensing and Control, 2004 IEEE International Conference on*, volume 1, pages 7–12, March 2004.

[Hsu 1999] Feng-hsiung Hsu. IBM's Deep Blue chess grandmaster chips. *IEEE Micro*, 19(2):70–81, 1999.

[Huang *et al.* 2008] Yueh-Min Huang, Yen-Hung Kuo, Yen-Ting Lin, and Shu-Chen Cheng. Toward interactive mobile synchronous learning environment with context-awareness service. *Computers & Education*, 51(3):1205–1226, 2008.

[Iwasawa *et al.* 1997] Shoichiro Iwasawa, Kazuyuki Ebihara, Jun Ohya, and Shigeo Morishima. Real-time estimation of human body posture from monocular thermal images. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 15–20. IEEE, 1997.

[Izard 1971] Carroll E. Izard. *The face of emotion*. East Norwalk, CT, US: Appleton-Century-Crofts, 1971.

[Izard 1994] Carroll E Izard. Innate and universal facial expressions: evidence from developmental and cross-cultural research. *Psychological Bulletin*, 115(2):288–299, Mar 1994.

[Jaimes and Sebe 2007] Alejandro Jaimes and Nicu Sebe. Multimodal human–computer interaction: A survey. *Computer vision and image understanding*, 108(1):116–134, 2007.

[James 1884] William James. What is an emotion? *Mind*, 9(34):188–205, 1884.

[Jeong *et al.* 2008] Jae-chan Jeong, Ho-chul Shin, and Dae-hwan Hwang. Real-time Upper Body Pose Detection using Stereo Vision ASIC. In *Proceedings of the 18th International Conference on Artificial Reality and Telexistence*, pages 238–241, 2008.

[Jia *et al.* 2014] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*, 2014.

[Jin *et al.* 2007] Shanshan Jin, So-Youn Park, and Ju-Jang Lee. Driver fatigue detection using a genetic algorithm. *Artificial Life and Robotics*, 11(1):87–90, 2007.

[Johns and Woolf 2006] Jeffrey Johns and Beverly Woolf. A dynamic mixture model to detect student motivation and proficiency. In *Proceedings of the National Conference on Artificial Intelligence*, volume 21, page 163. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.

[Johnstone and Scherer 2000] Tom Johnstone and KR Scherer. Vocal communication of emotion. *Handbook of Emotion*, pages 220–235, 2000.

[Jones *et al.* 2009] Geraldine Jones, Gabriele Edwards, and Alan Reid. How Can Mobile SMS Communication Support and Enhance a First Year Undergraduate Learning Environment?. *ALT-J: Research in Learning Technology*, 17(3):201–218, 2009.

[Juang *et al.* 2009] Chia-Feng Juang, Chia-Ming Chang, Jiuh-Rou Wu, and Demei Lee. Computer vision-based human body segmentation and posture estimation. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 39(1):119–133, 2009.

[Juslin and Scherer 2005] Patrick N Juslin and Klaus R Scherer. Vocal expression of affect. *The new handbook of methods in nonverbal behavior research*, pages 65–135, 2005.

[Kagan 1984] J. Kagan. *The Nature of the Child*. Gosset/Putnam Press, New York, NY, 1984.

[Kai *et al.* 2015] Shiming Kai, Luc Paquette, Ryan S Baker, Nigel Bosch, Sidney D'Mello, Jaclyn Ocumpaugh, Valerie Shute, and Matthew Ventura. A comparison of video-based and interaction-based affect detectors in physics playground. *International Educational Data Mining Society*, 2015.

[Kapoor and Picard 2001] Ashish Kapoor and Rosalind W. Picard. A Real-time Head Nod and Shake Detector. In *Proceedings of the 2001 Workshop on Perceptive User Interfaces*, PUI '01, pages 1–5, New York, NY, USA, 2001. ACM.

[Kapoor and Picard 2002] Ashish Kapoor and Rosalind W Picard. Real-time, fully automatic upper facial feature tracking. In *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*, pages 8–13. IEEE, 2002.

[Kapoor and Picard 2005] Ashish Kapoor and Rosalind W Picard. Multimodal affect recognition in learning environments. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 677–682. ACM, 2005.

[Kapoor *et al.* 2007] Ashish Kapoor, Winslow Burleson, and Rosalind W Picard. Automatic prediction of frustration. *International Journal of Human-Computer Studies*, 65(8):724–736, August 2007.

[Kiyoshi Kawaguchi 2000] Kiyoshi Kawaguchi. *Linear Separability and the XOR Problem*. http://www.ece.utep.edu/research/webfuzzy/docs/kk-thesis/kk-thesis-html/node19.html, 2000. Online; accessed 2017-01-20.

[Klein and Celik 2017a] Richard Klein and Turgay Celik. Engage: Live Self Reported Engagement for Large Classes. *IEEE Transactions on Learning Technologies*, 2017. Under review.

[Klein and Celik 2017b] Richard Klein and Turgay Celik. The Wits Intelligent Teaching System: Detecting Student Engagement During Lectures Using Convolutional Neural Networks. In *Image Processing, 2017. ICIP'17. 2017 International Conference on*. IEEE, 2017. Accepted.

[Klein and Celik 2017c] Richard Klein and Turgay Celik. Visualization of Audience Interest: An Interest Map for Reporting Live Audience Engagement. *IEEE Transactions on Learning Technologies*, 2017. Under review.

[Klein and Celik 2017d] Richard Klein and Turgay Celik. Wits Intelligent Teaching System: A video dataset and computer vision system for student action recognition in the classroom. *IEEE Transactions on Affective Computing*, 2017. Under review.

[Koenderink 1986] Jan J Koenderink. Optic flow. *Vision research*, 26(1):161–179, 1986.

[Kort *et al.* 2001] Barry Kort, Rob Reilly, and Rosalind W Picard. An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion. In *Advanced Learning Technologies, IEEE International Conference on*, pages 0043–0043. IEEE Computer Society, 2001.

[Kouloumpis *et al.* 2011] Efthymios Kouloumpis, Theresa Wilson, and Johanna D Moore. Twitter sentiment analysis: The good the bad and the omg! *Icwsm*, 11:538–541, 2011.

[Krizhevsky *et al.* 2012] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[Kroes 2007] Stefan Kroes. Detecting Boredom in Meetings. *Enschede, Netherlands, University of Twente*, pages 1–5, 2007.

[Kumar *et al.* 2009] Neeraj Kumar, Alexander C Berg, Peter N Belhumeur, and Shree K Nayar. Attribute and simile classifiers for face verification. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 365–372. IEEE, 2009.

[Kyriakides and Leventhall 1977] K Kyriakides and HG Leventhall. Some effects of infrasound on task performance. *Journal of Sound and Vibration*, 50(3):369–388, 1977.

[Lan and Sie 2010] Yu-Feng Lan and Yang-Siang Sie. Using RSS to support mobile learning based on media richness theory. *Computers & Education*, 55(2):723–732, 2010.

[Landis and Koch 1977] J Richard Landis and Gary G Koch. The measurement of observer agreement for categorical data. *biometrics*, pages 159–174, 1977.

[Landström and Byström 1984] U. Landström and M Byström. Infrasonic threshold levels of physiological effects effects. *Journal of Low Frequency Noise and Vibration*, 3:167–173, 1984.

[Landström and Byström 1985] U. Landström and M Byström. Changes in wakefulness during exposure to noise at 42Hz, 1000Hz and individual EEG frequencies. *Journal of Low Frequency Noise and Vibration*, 4:27–33, 1985.

[Landstrom *et al.* 1991] U Landstrom, A Kjellberg, L Söderberg, and B Nordström. The effects of broadband, tonal and masked ventilation noise on performance, wakefulness and annoyance. *Journal of low frequency noise & vibration*, 10(4):112–122, 1991.

[Lange 1885] Carl Georg Lange. The mechanism of the emotions. *The Classical Psychologists. Boston: Houghton Mifflin*, 1912, 1885.

[Langley 1996] Pat Langley. *Elements of machine learning*. Morgan Kaufmann, 1996.

[Larson and Richards 1991] Reed W Larson and Maryse H Richards. Boredom in the middle school years: Blaming schools versus blaming students. *American Journal of Education*, 1991.

12

[Lau *et al.* 2014] Rynson WH Lau, Neil Y Yen, Frederick Li, and Benjamin Wah. Recent development in multimedia e-learning technologies. *World Wide Web*, 17(2):189–198, 2014.

[LeCun *et al.* 1989] Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.

[Leventhall *et al.* 2003] Geoff Leventhall, Peter Pelmear, and Stephen Benton. *A review of published research on low frequency noise and its effects*. Technical report, Department for Environment, Food and Rural Affairs, 2003.

[Lienhart *et al.* 2003] Rainer Lienhart, Alexander Kuranov, and Vadim Pisarevsky. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. In *Joint Pattern Recognition Symposium*, pages 297–304. Springer, 2003.

[Lin *et al.* 2007] Hsuan-Tien Lin, Chih-Jen Lin, and Ruby C Weng. A note on Platt's probabilistic outputs for support vector machines. *Machine learning*, 68(3):267–276, 2007.

[Lindquist *et al.* 2006] Kristen A Lindquist, Lisa Feldman Barrett, Eliza Bliss-Moreau, and James A Russell. Language and the perception of emotion. *Emotion*, 6(1):125, 2006.

[Lindquist *et al.* 2007] David Lindquist, Tamara Denning, Michael Kelly, Roshni Malani, William G Griswold, and Beth Simon. Exploring the potential of mobile phones for active learning in the classroom. In *ACM SIGCSE Bulletin*, volume 39, pages 384–388. ACM, 2007.

[Liu *et al.* 2014] Shuaicheng Liu, Lu Yuan, Ping Tan, and Jian Sun. Steadyflow: Spatially smooth optical flow for video stabilization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4209–4216, 2014.

[Lowe and Laffey 2011] Ben Lowe and Des Laffey. Is Twitter for the birds? Using Twitter to enhance student learning in a marketing course. *Journal of Marketing Education*, page 0273475311410851, 2011.

[Lu and Viehland 2008] Xu Lu and Dennis Viehland. Factors influencing the adoption of mobile learning. *ACIS 2008 Proceedings*, page 56, 2008.

[Lucas *et al.* 1977] RW Lucas, PJ Mullin, CB Luna, and DC McInroy. Psychiatrists and a computer as interrogators of patients with alcohol-related illnesses: a comparison. *The British Journal of Psychiatry*, 131(2):160–167, 1977.

[Lucas *et al.* 1981] Bruce D Lucas, Takeo Kanade, et al. An iterative image registration technique with an application to stereo vision. 1981.

[MacHardy *et al.* 2012] Zachary M MacHardy, Kenneth Syharath, and Prasun Dewan. Engagement analysis through computer vision. In *Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2012 8th International Conference on*, pages 535–539. IEEE, 2012.

[Maclay *et al.* 1938] W.S. Maclay, E. Guttmann, and W. Mayer-Gross. Spontaneous Drawings as an Approach to some Problems of Psychopathology. *Proceedings of the Royal Society of Medicine*, 31(11):1337–1350, September 1938.

[Mallat 1990] SG Mallat. Multiresolution approach to wavelets in computer vision. In *Wavelets*, pages 313–327. Springer, 1990.

[Manning *et al.* 2008] Christopher D Manning, Prabhakar Raghavan, Hinrich Schütze, et al. *Introduction to information retrieval*, volume 1. Cambridge university press Cambridge, 2008.

[Markett *et al.* 2006] Carina Markett, I Arnedillo Sánchez, Stefan Weber, and Brendan Tangney. Using short message service to encourage interactivity in the classroom. *Computers & Education*, 46(3):280–293, 2006.

[Marsland 2015] Stephen Marsland. *Machine learning: an algorithmic perspective*. CRC press, 2015.

[McCulloch and Pitts 1943] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[McDaniel *et al.* 2007] B. McDaniel, Sidney D'Mello, B. King, P. Chipman, K. Tapp, and A. Graesser. Facial Features for Affective State Detection in Learning Environments. In *Proceedings of the 29th. Annual Meeting of the Cognitive Science Society*, 2007.

[McQuiggan and Lester 2006] Scott W McQuiggan and James C Lester. Diagnosing self-efficacy in intelligent tutoring systems: An empirical study. In *Intelligent Tutoring Systems*, pages 565–574. Springer, 2006.

[Mehrabian 1977] Albert Mehrabian. *Nonverbal communication*. Transaction Publishers, 1977.

[Mendell and Heath 2005] M. J. Mendell and G. A. Heath. Do indoor pollutants and thermal conditions in schools influence student performance? A critical review of the literature. *Indoor Air*, 15(1):27–52, 2005.

[Merla and Romani 2007] A. Merla and G.L. Romani. Thermal Signatures of Emotional Arousal: A Functional Infrared Imaging Study. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 247–249, Aug 2007.

[Mikulas and Vodanovich 1993] William L Mikulas and Stephen J Vodanovich. The essence of boredom. *The Psychological Record*, 1993.

[Minsky 1961] M. Minsky. Steps toward artificial intelligence. *Proceedings of the IRE*, 49(1):8–30, Jan 1961.

[Minsky 2007] Marvin Minsky. *The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind*. Simon and Schuster, 2007.

[Moeslund *et al.* 2006] Thomas B Moeslund, Adrian Hilton, and Volker Krüger. A survey of advances in vision-based human motion capture and analysis. *Computer vision and image understanding*, 104(2):90–126, 2006.

[Mohri *et al.* 2012] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2012.

[Montepare *et al.* 1999] Joann Montepare, Elissa Koff, Deborah Zaitchik, and Marilyn Albert. The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behavior*, 23(2):133–152, 1999.

[Morency *et al.* 2007] Louis-Philippe Morency, Candace Sidner, Christopher Lee, and Trevor Darrell. Head gestures for perceptual interfaces: The role of context in improving recognition. *Artificial Intelligence*, 171(8-9):568–585, 2007.

[Mota and Picard 2003] Selene Mota and Rosalind W Picard. Automated posture analysis for detecting learner's interest level. In *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Conference on*, volume 5, pages 49–49. IEEE, 2003.

[Motiwalla 2007] Luvai F Motiwalla. Mobile learning: A framework and evaluation. *Computers & education*, 49(3):581–596, 2007.

[Murphy-Chutorian and Trivedi 2009] Erik Murphy-Chutorian and Mohan M Trivedi. Head pose estimation in computer vision: A survey. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(4):607–626, 2009.

[Nickolls *et al.* 2008] John Nickolls, Ian Buck, Michael Garland, and Kevin Skadron. Scalable parallel programming with cuda. *Queue*, 6(2):40–53, March 2008.

[Nicol and Boyle 2003] David J Nicol and James T Boyle. Peer instruction versus class-wide discussion in large classes: a comparison of two interaction methods in the wired classroom. *Studies in Higher Education*, 28(4):457–473, 2003.

[NVIDIA 2015] NVIDIA. *NVIDIA TITAN X Graphics Card with Pascal*. https://www.nvidia.com/en-us/geforce/products/10series/titan-x-pascal/, 2015. Online; accessed 2017-02-18.

[Ocumpaugh 2012] Jaclyn Ocumpaugh. Baker-Rodrigo observation method protocol (BROMP) 1.0. Training manual version 1.0. *Baker-Rodrigo Observation Method Protocol (BROMP) 1.0. Training Manual version 1.0, Tech. Rep*, 2012.

[Öhman and Soares 1998] Arne Öhman and Joaquim JF Soares. Emotional conditioning to masked stimuli: expectancies for aversive outcomes following nonrecognized fear-relevant stimuli. *Journal of Experimental Psychology: General*, 127(1):69, 1998.

[Ojala and Pietikäinen 1999] Timo Ojala and Matti Pietikäinen. Unsupervised texture segmentation using feature distributions. *Pattern Recognition*, 32(3):477–486, 1999.

[Ojala *et al.* 1996] Timo Ojala, Matti Pietikinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):5159, 1996.

15

[Ojala *et al.* 2002] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution grayscale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7):971–987, 2002.

[OpenCV 2014] OpenCV. *Introduction to Support Vector Machines*. http://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html, 2014. Online; accessed 2017-01-20.

[OpenCV 2015] OpenCV. *Optical Flow*. http://docs.opencv.org/3.2.0/d7/d8b/tutorial_py_lucas_kanade.html, 2015. Online; accessed 2017-01-20.

[Ortony and Turner 1990] Andrew Ortony and Terence J Turner. What's basic about basic emotions? *Psychological review*, 97(3):315, 1990.

[Ortony 1990] Andrew Ortony. *The cognitive structure of emotions*. Cambridge university press, 1990.

[Pak and Paroubek 2010] Alexander Pak and Patrick Paroubek. Twitter as a corpus for sentiment analysis and opinion mining. In *LREc*, volume 10, pages 1320–1326, 2010.

[Palloff and Pratt 1999] R.M. Palloff and K. Pratt. *Building learning communities in cyberspace: effective strategies for the online classroom*. Jossey-Bass, San Francisco, 1999.

[Panksepp 1998] Jaak Panksepp. *Affective neuroscience: The foundations of human and animal emotions*. Oxford university press, 1998.

[Pantic and Patras 2006] Maja Pantic and Ioannis Patras. Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 36(2):433–449, 2006.

[Pantic and Rothkrantz 2003] Maja Pantic and Leon JM Rothkrantz. Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE*, 91(9):1370–1390, 2003.

[Paquette *et al.* 2016] Luc Paquette, Jonathan Rowe, Ryan Baker, Bradford Mott, James Lester, Jeanine DeFalco, Keith Brawner, Robert Sottilare, and Vasiliki Georgoulas. Sensor-free or sensor-full: A comparison of data modalities in multi-channel affect detection. *International Educational Data Mining Society*, 2016.

[Patten *et al.* 2006] Bryan Patten, Inmaculada Arnedillo Sánchez, and Brendan Tangney. Designing collaborative, constructionist and contextual applications for handheld devices. *Computers & education*, 46(3):294–308, 2006.

[Persson Waye and Rylander 2001] K Persson Waye and R Rylander. The prevalence of annoyance and effects after long-term exposure to low-frequency noise. *Journal of sound and vibration*, 240(3):483–497, 2001.

[Persson Waye *et al.* 1997] Kerstin Persson Waye, R Rylander, S Benton, and HG Leventhall. Effects on performance and work quality due to low frequency ventilation noise. *Journal of Sound and Vibration*, 205(4):467–474, 1997.

[Phillips *et al.* 2011] P Jonathon Phillips, J Ross Beveridge, Bruce A Draper, Geof Givens, Alice J O'Toole, David S Bolme, Joseph Dunlop, Yui Man Lui, Hassan Sahibzada, and Samuel Weimer. An introduction to the good, the bad, & the ugly face recognition challenge problem. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, pages 346–353. IEEE, 2011.

[Picard 1995] Rosalind W. Picard. *Affective Computing*. Technical report, MIT Media Laboratory, 1995. http://www.media.mit.edu/ picard/.

[Picard 1997] Rosalind W. Picard. *Affective Computing*. MIT Press, 1997.

[Picard 2003a] Rosalind W. Picard. Affective computing: challenges. *International Journal of Human-Computer Studies*, 59(1–2):55–64, 2003. Applications of Affective Computing in Human-Computer Interaction.

[Picard 2003b] Rosalind W Picard. What does it mean for a computer to "have" emotions. *Emotions in humans and artifacts*, pages 87–102, 2003.

[Platt and others 1999] John Platt et al. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, 10(3):61–74, 1999.

[Poulsen *et al.* 2017] Andreas Trier Poulsen, Simon Kamronn, Jacek Dmochowski, Lucas C Parra, and Lars Kai Hansen. EEG in the classroom: Synchronised neural recordings during video presentation. *Scientific Reports*, 7, 2017.

[Puri *et al.* 2005] Colin Puri, Leslie Olson, Ioannis Pavlidis, James Levine, and Justin Starren. StressCam: Non-contact Measurement of Users' Emotional States Through Thermal Imaging. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1725–1728, Portland, OR, USA, 2005. ACM.

[Raca and Dillenbourg 2013] Mirko Raca and Pierre Dillenbourg. System for assessing classroom attention. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 265–269. ACM, 2013.

[Raca *et al.* 2015] Mirko Raca, Lukasz Kidzinski, and Pierre Dillenbourg. Translating head motion into attention-towards processing of student's body-language. In *Proceedings of the 8th International Conference on Educational Data Mining*, number EPFL-CONF-207803, 2015.

[Ramanan 2008] Deva Ramanan. *ICS 273A Machine Learning, Lecture 11*. http://www.ics.uci.edu/~dramanan/teaching/ics273a_winter08/lectures/lecture11.pdf, 2008. Online; accessed 2017-01-26.

[Rau *et al.* 2008] Pei-Luen Patrick Rau, Qin Gao, and Li-Mei Wu. Using mobile communication technology in high school education: Motivation, pressure, and learning performance. *Computers & Education*, 50(1):1–22, 2008.

[Reisenzein *et al.* 2013] Rainer Reisenzein, Markus Studtmann, and Gernot Horstmann. Coherence between emotion and facial expression: Evidence from laboratory experiments. *Emotion Review*, 5(1):16–23, 2013.

[Richardson 2004] Iain E Richardson. *H. 264 and MPEG-4 video compression: video coding for next-generation multimedia.* John Wiley & Sons, 2004.

[Robinson and West 1992] Rachael Robinson and Robert West. A comparison of computer and questionnaire methods of history-taking in a genito-urinary clinic. *Psychology and Health*, 6(1-2):77–84, 1992.

[Roseman *et al.* 1990] Ira J Roseman, Martin S Spindel, and Paul E Jose. Appraisals of emotion-eliciting events: Testing a theory of discrete emotions. *Journal of Personality and Social Psychology*, 59(5):899, 1990.

[Roseman 1984] Ira J Roseman. Cognitive determinants of emotion: A structural theory. *Review of Personality & Social Psychology*, 1984.

[Rosenfeld 1988] Azriel Rosenfeld. Computer vision: basic principles. *Proceedings of the IEEE*, 76(8):863–868, 1988.

[Royer and Walles 2007] James M Royer and Rena Walles. Influences of gender, motivation and socioeconomic status on mathematics performance. *Why is math so hard for some children. Baltimore, MD: Paul H. Brookes Publishing Co*, pages 349–368, 2007.

[Russell *et al.* 2003a] James A Russell, Jo-Anne Bachorowski, and José-Miguel Fernández-Dols. Facial and vocal expressions of emotion. *Annual review of psychology*, 54(1):329–349, 2003.

[Russell *et al.* 2003b] James A Russell, Jo-Anne Bachorowski, and José-Miguel Fernández-Dols. Facial and vocal expressions of emotion. *Annual review of psychology*, 54(1):329–349, 2003.

[Russell 1994] James A Russell. Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological bulletin*, 115(1):102, 1994.

[Russell 2003] James A Russell. Core affect and the psychological construction of emotion. *Psychological review*, 110(1):145, 2003.

[Schachter and Singer 1962] Stanley Schachter and Jerome Singer. Cognitive, social, and physiological determinants of emotional state. *Psychological review*, 69(5):379, 1962.

[Schaeffer *et al.* 2007] Jonathan Schaeffer, Neil Burch, Yngvi Björnsson, Akihiro Kishimoto, Martin Müller, Robert Lake, Paul Lu, and Steve Sutphen. Checkers is solved. *science*, 317(5844):1518–1522, 2007.

[Scherer and Ellgring 2007a] Klaus R Scherer and Heiner Ellgring. Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion*, 7(1):113, 2007.

[Scherer and Ellgring 2007b] Klaus R Scherer and Heiner Ellgring. Multimodal expression of emotion: Affect programs or componential appraisal patterns? *Emotion*, 7(1):158, 2007.

[Scherer *et al.* 2001] Klaus R Scherer, Angela Ed Schorr, and Tom Ed Johnstone. *Appraisal processes in emotion: Theory, methods, research.* Oxford University Press, 2001.

[Schneider and Josephs 1991] Klaus Schneider and Ingrid Josephs. The expressive and communicative functions of preschool children's smiles in an achievement-situation. *Journal of nonverbal behavior*, 15(3):185–198, 1991.

[Scott 2011] G.D. Scott. Doodling and the default network of the brain. *The Lancet*, 378(9797):1133 – 1134, September 2011.

[Sharma *et al.* 1998] Rajeev Sharma, Vladimir I Pavlovic, and Thomas S Huang. Toward multimodal human-computer interface. *Proceedings of the IEEE*, 86(5):853–869, 1998.

[Sharples and others 2006] Mike Sharples et al. Big issues in mobile learning. In *Report of a workshop by the Kaleidoscope Network of Excellence Mobile Learning Initiative*. LSRI, University of Nottingham Nottingham, 2006.

[Sharples *et al.* 2005] Mike Sharples, Josie Taylor, and Giasemi Vavoula. Towards a theory of mobile learning. In *Proceedings of mLearn*, volume 1, pages 1–9, 2005.

[Shastri *et al.* 2009] D. Shastri, A. Merla, P. Tsiamyrtzis, and I. Pavlidis. Imaging Facial Signs of Neurophysiological Responses. *Biomedical Engineering, IEEE Transactions on*, 56(2):477–484, Feb 2009.

[Shavlik and Dietterich 1990] Jude W Shavlik and Thomas Glen Dietterich. *Readings in machine learning*. Morgan Kaufmann, 1990.

[Shendell *et al.* 2004] D.G. Shendell, R. Prill, W.J. Fisk, M.G. Apte, D. Blake, and D. Faulkner. Associations between classroom CO_2 concentrations and student attendance in Washington and Idaho. *Indoor Air*, 14(5):333–341, 2004.

[Shernoff *et al.* 2003] David J Shernoff, Mihaly Csikszentmihalyi, Barbara Shneider, and Elisa Steele Shernoff. Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, 18(2):158, 2003.

[Shute *et al.* 2015] Valerie J Shute, Sidney D'Mello, Ryan Baker, Kyunghwa Cho, Nigel Bosch, Jaclyn Ocumpaugh, Matthew Ventura, and Victoria Almeda. Modeling how incoming knowledge, persistence, affective states, and in-game progress influence student learning from an educational game. *Computers & Education*, 86:224–235, 2015.

[Silver *et al.* 2016] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, van den Driessche George, Julian Schrittwieser, Loannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, jan 2016.

[Simon 1983] Herbert A Simon. Why should machines learn? In *Machine learning*, pages 25–37. Springer, 1983.

[Smith and Ellsworth 1985] Craig A Smith and Phoebe C Ellsworth. Patterns of cognitive appraisal in emotion. *Journal of personality and social psychology*, 48(4):813, 1985.

[Smith *et al.* 2008] Kevin Smith, Sileye O Ba, Jean-Marc Odobez, and Daniel Gatica-Perez. Tracking the visual focus of attention for a varying number of wandering people. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(7):1212–1229, 2008.

[South African Reserve Bank 2017] South African Reserve Bank. *Banknotes*. http://banknotes.resbank.co.za/banknotes, 2017. Online; accessed 2017-01-20.

[Stets and Turner 2008] Jan E Stets and Jonathan H Turner. The sociology of emotions. *Handbook of emotions*, pages 32–46, 2008.

[Strang *et al.* 1993] Gilbert Strang, Gilbert Strang, Gilbert Strang, and Gilbert Strang. *Introduction to linear algebra*, volume 3. Wellesley-Cambridge Press Wellesley, MA, 1993.

[Sun *et al.* 2005] Nanfei Sun, M. Garbey, A. Merla, and I. Pavlidis. Imaging the cardiovascular pulse. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 416–421 vol. 2, June 2005.

[Sutton and Barto 1998] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

[Szeliski 2010] Richard Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

[Taigman *et al.* 2014] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1701–1708, 2014.

[Tarhini *et al.* 2015] Ali Tarhini, Mohammad Hassouna, Muhammad Sharif Abbasi, and Jorge Orozco. Towards the Acceptance of RSS to Support Learning: An Empirical Study to Validate the Technology Acceptance Model in Lebanon. *Electronic Journal of e-Learning*, 13(1):30–41, 2015.

[The Qt Company 2016] The Qt Company. *Qt – Cross-platform software development for embedded & desktop*. https://www.qt.io/, 2016. Online; accessed 2016-12-12.

[Tomkins 1962] Silvan S Tomkins. *Affect, imagery, consciousness: Vol. I. The positive affects.* Springer, 1962.

[Turing 1950] Alan. M. Turing. Computing Machinery and Intelligence. *Mind*, LIX(236):433–460, 1950.

[Turk and Pentland 1991] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. In *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 586–591, Jun 1991.

[van de Pol *et al.* 2011] Janneke van de Pol, Monique Volman, and Jos Beishuizen. Patterns of contingent teaching in teacher–student interaction. *Learning and Instruction*, 21(1):46–57, 2011.

[Van den Bergh *et al.* 2008] Michael Van den Bergh, Esther Koller-Meier, and Luc Van Gool. Fast body posture estimation using volumetric features. In *Motion and video Computing, 2008. WMVC 2008. IEEE Workshop on*, pages 1–8. IEEE, 2008.

[Vapnik 1995] V Vapnik. *The nature of statistical learning*, 1995.

[Viola and Jones 2001] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–511. IEEE, 2001.

[Viola and Jones 2004] Paul Viola and Michael J Jones. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004.

[Vlachos *et al.* 2002] Michail Vlachos, Carlotta Domeniconi, Dimitrios Gunopulos, George Kollios, and Nick Koudas. Non-linear dimensionality reduction techniques for classification and visualization. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '02, pages 645–651, New York, NY, USA, 2002. ACM.

[Vonderwell 2003] S Vonderwell. An examination of asynchronous communication experiences and perspectives of students in an online course: a case study. *The Internet and Higher Education*, 6(1):77–90, 2003.

[Walk and Walters 1988] RD Walk and KL Walters. Perception of the smile and other emotions of the body and face at different distances. In *Bulletin of the Psychonomic Society*, volume 26, pages 510–510. Psychonomic Soc Int 1710 Fortview Rd, Austin, TX 78704, 1988.

[Wallbott 1998] Harald G Wallbott. Bodily expression of emotion. *European journal of social psychology*, 28(6):879–896, 1998.

[Wang and Shi 2005] Tiesheng Wang and Pengfei Shi. Yawning detection for determining driver drowsiness. In *VLSI Design and Video Technology, 2005. Proceedings of 2005 IEEE International Workshop on*, pages 373–376, May 2005.

[Wang *et al.* 2006] Qiong Wang, Jingyu Yang, Mingwu Ren, and Yujie Zheng. Driver Fatigue Detection: A Survey. In *Intelligent Control and Automation, 2006. WCICA 2006. The Sixth World Congress on*, volume 2, pages 8587–8591, 2006.

[Wang *et al.* 2011] H. Wang, A. Klser, C. Schmid, and C. L. Liu. Action recognition by dense trajectories. In *CVPR 2011*, pages 3169–3176, June 2011.

[Weinberg 2012] Zack Weinberg. *Svm separating hyperplanes (SVG).* https://en.wikipedia.org/wiki/File:Svm_separating_hyperplanes_(SVG).svg, 2012. Online; accessed 2017-01-20.

[Wentzel and Asher 1995] Kathryn R Wentzel and Steven R Asher. The academic lives of neglected, rejected, popular, and controversial children. *Child development*, 66(3):754–763, 1995.

[Whitehill *et al.* 2014] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J.R. Movellan. The Faces of Engagement: Automatic Recognition of Student Engagementfrom Facial Expressions. *IEEE Transactions on Affective Computing*, 5(1):86–98, March 2014.

[Witten *et al.* 2016] Ian H Witten, Eibe Frank, Mark A Hall, and Christopher J Pal. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.

[Woolf *et al.* 2009] Beverly Woolf, Winslow Burleson, Ivon Arroyo, Toby Dragon, David Cooper, and Rosalind Picard. Affect-aware tutors: recognising and responding to student affect. *International Journal of Learning Technology*, 4(3):129–164, 2009.

[Wu *et al.* 2012] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John V Guttag, Frédo Durand, and William T Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph.*, 31(4):65, 2012.

[Yeasin *et al.* 2006] M. Yeasin, B. Bullot, and R. Sharma. Recognition of facial expressions and measurement of levels of interest from video. *IEEE Transactions on Multimedia*, 8(3):500–508, June 2006.

[Zeng *et al.* 2009] Zhihong Zeng, Maja Pantic, Glenn I Roisman, and Thomas S Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(1):39–58, 2009.

[Zimmerman 2000] Barry J Zimmerman. Self-efficacy: An essential motive to learn. *Contemporary educational psychology*, 25(1):82–91, 2000.