

Involuntary Gesture Recognition for Predicting Cerebral Palsy in High-Risk Infants

Mohan Singh
School of Computer Science and Informatics
University College Dublin
Dublin 4, Ireland
{mohan.singh}@ucd.ie

Donald J. Patterson
Department of Informatics
University of California, Irvine
Irvine, CA, USA 92697-3440
djp3@ics.uci.edu

Abstract

*In this paper we describe a system that leverages accelerometers to recognize a particular involuntary gesture in babies that have been born preterm. These gestures, known as cramped-synchronized general movements are highly correlated with a diagnosis of Cerebral Palsy. In order to test our system we recorded data from 10 babies admitted to the newborn intensive care unit at the UCI Medical Center. We applied machine learning techniques to features based on their data and were able to obtain accuracies between 70% and 90% depending on the relative cost of false positives and false negatives. Validated video observation annotations were utilized as ground truth. Finally, we conducted an analysis to understand the basis of the algorithmic predictions.*¹

1. Background

Over the last three decades, in industrialized countries, there has been an increase in the number of children born pre-term (less than 37 weeks gestational age). In the U.S. the number has risen 34% from 9.5% in 1981 to 12.7% in 2005 and continues to rise [12]. The survival rates of very early preterm babies have increased due to technological advances and improved medical specialist collaboration.

However, these babies are at increased risk for neurological problems such as Cerebral Palsy, mental retardation and sensory impairments. Cerebral Palsy has been shown to occur in 7.2% of babies born before 26 weeks of gestational age [24].

The term “Cerebral Palsy” (CP) refers to a number of neurological disorders that appear in infancy or early childhood and affect body movements and coordination permanently but which aren’t degenerative. It is caused by abnor-

malities inside the brain that disrupt the brain’s ability to control movement and posture [19]. The early diagnosis of CP and other motor abnormalities would enable clinicians to implement early interventions that may provide an improved quality of life for the patient. Unfortunately, early diagnosis is difficult and diagnosis is usually made between several months to 2 years of age [2].

Current medical practice to diagnosis CP is for the physician to conduct a neurological exam. However in a study of over 40,000 children conducted in 1992, only 23% of infants with CP (N=128) had an abnormal neurological exam in infancy [18]. In practice, doctors end up diagnosing cerebral palsy by evaluating a child’s motor skills and medical history over several years. Neuroimaging techniques, such as cranial ultrasound, CT scans and MRI scans can also assist in diagnosis.

However, other less clinically practical examinations have been developed to diagnose CP. For example, Prechtl et. al. developed an observational technique that evaluates the motor quality of an infant’s movement from one-hour video sessions. The two key markers are involuntary gestures that are qualitatively identified as fidgety movements and cramped-synchronized general movements (CSGM). CSGMs look rigid and are characterized by all limb and trunk muscles contracting and relaxing almost simultaneously. The lack of the former and the presence of the latter are indicative of CP. Persistent CSGMs have been shown to predict CP with a specificity of 93% in preterm babies [11]

Unfortunately this technique is not clinically practical because it requires an hour of video annotation by a trained specialist. Cost, observer fatigue and lack of trained specialists all contribute to making this approach infeasible. We hypothesized that this assessment would be highly amenable to technological intervention building on work in the wearable and ubiquitous computing community on gesture recognition.

¹this work partially supported by Science Foundation Ireland

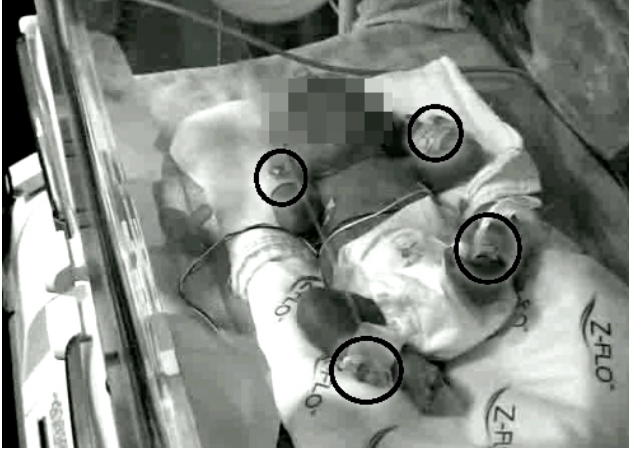


Figure 1. Baby in the NICU with accelerometers on each limb.

Developments in gesture recognition from worn sensors have been rapidly progressing in recent years as a result of increased processing power on mobile devices and increasing availability of complex sensors [29], software [14] and algorithms [16, 13]. Although there is significant overlap, gesture recognition tends to be separated from activity recognition by whether the goal is to use sensor measurements to *label or log* an activity or to *explicitly control or inform* a system with a motion. Activity recognition tends toward the former while gesture recognition tends toward the latter. Additionally activity recognition focuses more on whole body motion through time and space, whereas gesture recognition tends to focus more on hand and limb motion. So, activity recognition strives to recognize Activities of Daily Living (ADLs) like “setting the table” [22] whereas gesture recognition tries to recognize the components of sign language (fenemes) [7]. Using this taxonomy, CSGM recognition leans more toward gesture recognition.

There is a large literature regarding body-worn gesture recognition systems. With regard to human-computer interaction, for example, forearm electromyography has been used to detect finger gestures [25], fabric-embedded sensors have been used to detect gross body motions [10, 9], wrist-worn accelerometers have been used to complete questionnaires [4] and for end-user gesture programming [5]. Worn-video cameras are increasingly being used for gesture recognition as well [17].

Specifically in relation to medically motivated gestures, researchers have detected characteristic motions of Parkinson’s disease [6, 21], used accelerometers to detect sleep/wake cycles in infants [26], detected eating gestures for dietary management [3], and detected stimming movements in autistic children [1, 15]

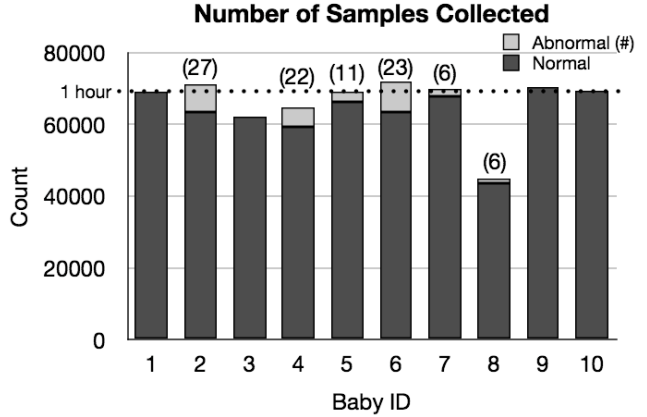


Figure 2. The number of normal and abnormal samples collected for each baby. The target number of samples was 68,400 per baby – an hour of data collected at 19Hz. The number of discrete CSGMs is shown in parentheses. Babies 1, 3, 9, and 10 exhibited no CSGMs. Baby 8 experienced a large amount of transmission interference.

1.1. Contribution

In this paper we evaluate the hypothesis that it is possible to recognize CSGMs by applying gesture recognition techniques to worn accelerometers. Such a system would be predictive of CP. In this paper we demonstrate one approach toward proving this hypothesis true and attempt to understand the underlying factors that affect its performance.

2. Methodology

In conjunction with the newborn intensive care unit (NICU) at the UCI Medical Center, we recruited 10 preterm babies with a gestational age at birth of between 23 and 36 weeks (see Figure 1). We recruited high-risk babies that had cerebral ultrasound abnormalities and low birth weight, both of which increase risk for CP. We videotaped each baby for 1 hour when they were between 30–43 weeks gestational age while wearing only a diaper in a temperature controlled isolette.

While videotaping, the babies were each wearing 4 custom-built accelerometers, two on the wrists and two on the ankles. Our accelerometer hardware, called an “Eco”, was custom designed by Dr. Pai Chou at the University of California, Irvine [20].

The Eco can measure acceleration along three axes from -3g to +3g and is light enough to be placed on a premature baby (2g) to measure changes in movement. Although not

utilized in this study, temperature and light sensing is also available on an Eco. The device is wireless and transmits its signal up to 10m on the same band as Bluetooth but using considerably less power. One computer can receive the input from up to 50 devices and the standard battery lasts 1.5 hours while sensing at full capacity.

As the video recording was being conducted, the accelerometers sampled data at approximately 19 Hz and wirelessly transmitted the data to a laptop located near by. The sampling was non-uniform as a byproduct of the round-robin implementation of the base station polling algorithm that continuously requests data from each accelerometer in turn. Subsequent packet loss and collision cause variable delays.

After the data collection session, the video data was transferred to a nurse trained in identifying CSGMs. The nurse annotated the video with start and stop time for each observed CSGM. The annotations were a 1-bit signal indicating the presence or absence of the CSGM and provided the gold standard that we used for our modeling.

3. Data Cleaning

Each baby generated approximately 70,000 data samples. Each sample consisted of a time stamp, T , and 12 real numbered features corresponding to 3 axes, x , y , and z of acceleration data from four limbs. Data was also collected from a fifth accelerometer located on the forehead, but proved to have negligible predictive power. By associating the nurse video annotations with the timestamps, each sample was given a classification, c , of “abnormal” or “normal”, corresponding to the presence or absence of a CSGM.

$$s_i = (T, x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3, x_4, y_4, z_4, c)$$

$$-3.0g \leq x, y, z \leq 3.0g$$

$$\{1, 2, 3, 4\} = \{\text{left arm, right arm, left leg, right leg}\}$$

Some of the data was invalidated as a result of transmission errors and due to interventions by NICU nursing staff. The former were automatically detected and removed from the data. The latter were removed from the data set by manually reviewing the recorded video. Examples of situations like this included pacifier adjustments by the NICU nurses, adjustments made for the comfort of the babies, and adjustments to the monitoring equipment. General characteristics of the data set are shown in Figure 2. A total of 95 CSGMs were identified in six of the ten babies.

Gravity exerts a constant 1.0g acceleration on each accelerometer. However, we do not maintain the pose of the baby in any of our data collection, so identifying the component of the acceleration that is due to a change in velocity vs. calibration drift vs. gravity is difficult. This is particularly true as the force due to gravity is generally split across

the 3-axes based on changing limb posture. To correct for systematic offsets, like these, that are constant over short time intervals, we subtracted the mean of the surrounding 10-second window from each sample, where the mean for each feature was calculated independently.

$$s'_i = s_i - \mu_{s_i \pm 10sec}$$

As a side effect of the non-uniform data rate, transmission errors, and contaminated data, the number of data points over which the mean was calculated varied for each data point. We intentionally limited the mean to this *temporal* window because the systematic problems that we sought to address varied relatively slowly over time regardless of the number of samples successfully collected.

The corrected sample, s'_i , was then used as the basis for detecting the magnitude of the acceleration observed for each of four limbs (left arm, right arm, left leg, right leg, respectively).

$$m_j = \sqrt{x_j'^2 + y_j'^2 + z_j'^2}$$

$$s''_i = (T, m_1, m_2, m_3, m_4, c)$$

4. Feature Creation

Based on the qualitative description of a CSGM we added 6 additional computed features to each sample: the maximum observed acceleration magnitude for the arms, the legs and the whole body and the product of the same:

$$\max(m_1, m_2), \max(m_3, m_4), \max(m_1, m_2, m_3, m_4)$$

$$(m_1 \times m_2), (m_3 \times m_4), (m_1 \times m_2 \times m_3 \times m_4)$$

For these 10 features we then computed a mean, max, min, standard deviation and z-score across a 2 second window centered on the current sample. Finally we added the Pearson correlation coefficient of the left arm with the right arm, and the left leg with the right leg, also aggregated over the same temporal window. The result was a vector with a time stamp, 52 features and a classification.

5. Cramped-Synchronized General Movement Modeling

5.1. Data Exploration

Using these features we applied three statistical machine learning techniques to predict the binary class using 10-fold cross validation: Naive Bayes, Support Vector Machines, and a pruned Decision Tree. As an additional baseline technique we predicted the most popular class, “normal”. While

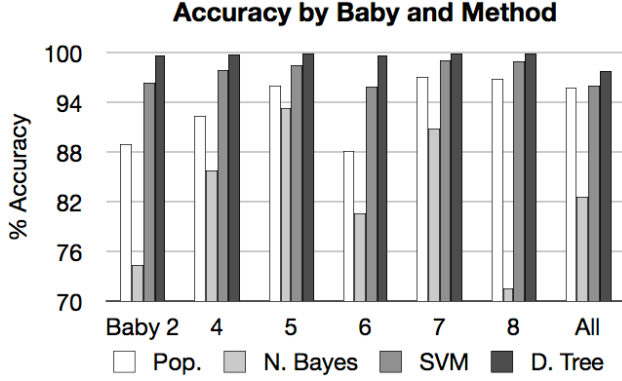


Figure 3. These are the accuracies obtained from different machine learning techniques predicting sample-by-sample abnormalities for the babies that exhibited CSGMs. “Pop.” always predicted the most popular class (“normal”). For each baby the results were trained and tested on a single baby at a time using cross-fold validation. The “All” column reflects the result from mixing all the baby’s data together prior to stratifying and modeling.

we were exploring the data, we trained the models in two ways. The first was on a baby-by-baby basis. In this case each baby’s data was treated as a separate data set and a cross-validated model was learned for each baby independently. Since only 6 babies exhibited CSGMs, statistics were only collected on those babies. In the second grouping we mixed all of the baby data together and trained a cross-validated model on the mixed data. While neither of these approaches are representative of how this technique would be applied clinically and are prone to high overfitting, they do provide insight into the variability of the data among babies. The results are shown in Figure 3.

The results from this experiment demonstrated that decision trees achieve the highest accuracy of the techniques tried (99.46%) followed by SVM’s (90.46%). Naive Bayes performed worse (70.43%) than just predicting the majority class (88.9%), but predicted true positives. All differences were statistically significant at ($p < 0.01$).

Understanding the reasons for these results is challenging. The relatively low performance of Naive Bayes suggests that this classification task depends heavily on the values of multiple features interacting with each other. With sufficient data, both SVMs and decision trees are able to learn functions of correlated features. As these results are based on analyses of continuous real numbered features, we were surprised to see decision trees outperformed SVMs.

In order to understand these results we isolated the most

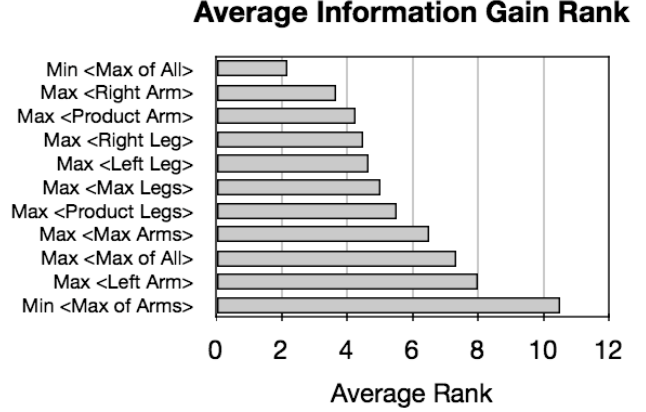
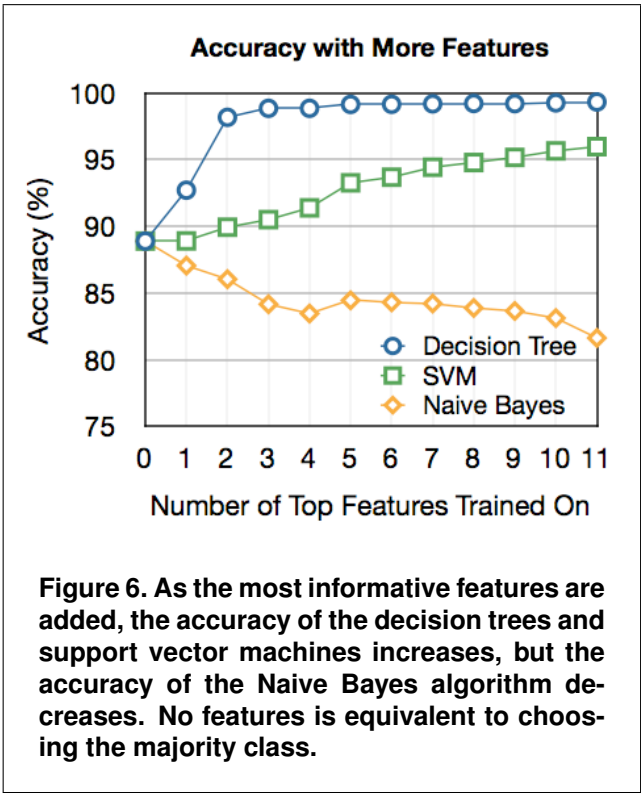
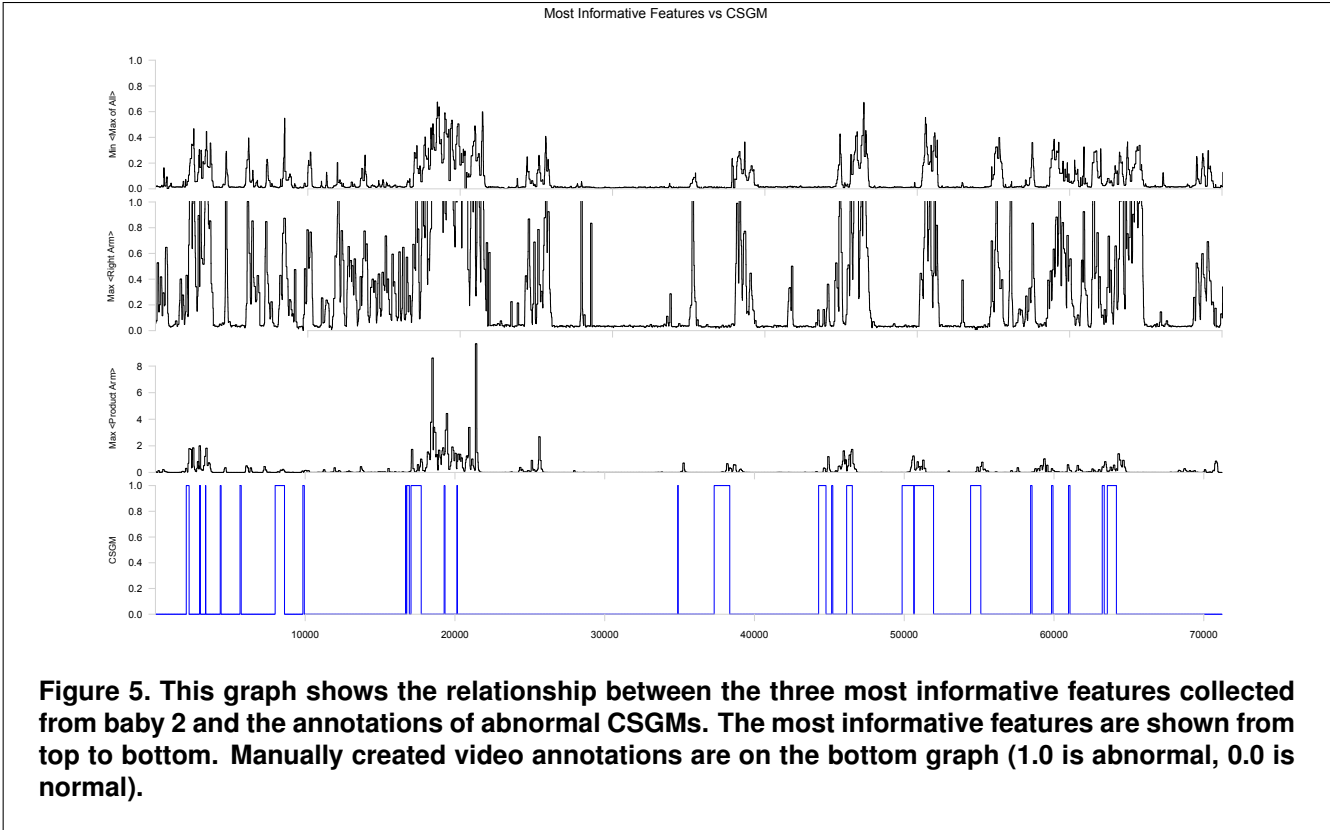


Figure 4. For each baby the information gain of each feature was ranked and the ranks were averaged across all babies and plotted above. The minimum across a 2 second window of the maximum acceleration magnitude of all limbs was the most informative feature.

informative features using an information gain analysis. For each baby we ranked each feature by its information gain independently of the others and then averaged the rank of that feature across all babies. Figure 4 shows the ranked results and Figure 5 shows a portion of the data from baby 2 for the top three most informative features. The most informative feature was the minimum across a 2 second window of the maximum acceleration magnitude of all limbs for a given sample. This points to the importance of CSGMs involving a sustained motion. CSGMs are not particularly high-energy motions however, and this possibly suggests that normal motions are bursty and will not sustain a continuously observed acceleration on all limbs for an entire 2 second window. The next two most informative features were based on recognizing high arm energy in a 2 second window followed by three features recognizing high leg energy. This suggests that the sensor signature of CSGMs is indeed a whole body experience and that perhaps arm motions are more rare in normal baby movements.

Aside from our intuitions, the role these top features play in the modeling is not clear. Figure 6 shows the increase in accuracy that is obtained by incrementally adding the most informative features to the data set for each algorithm. When using decision trees the first feature improves accuracy from a baseline of 88.90% to 92.72%, a decrease in error of 34%. The second feature improves accuracy to 98.18%, another decrease in error of 75%. Eight of the next nine features continue to improve accuracy, suggesting that there is significant additional information in each of these features.



Decision trees show a much more dramatic increase with the addition of the first two features. This is in contrast to a steady improvement in SVMs and a trade-off in decreased accuracy for an improved true positive rate for the naive bayes algorithm. Since the most informative feature was an aggregate measure of all four sensors, these results do not provide evidence for monitoring a subset of limbs with less sensors.

An analysis of the decision trees that are created is less than illuminating. Rather than simply making a thresholding decision, the decision tree that is formed *based on one feature* contained 899 nodes and 450 leaves in the pruned tree. As we increased the available features, the decision tree grew increasingly complicated and performed better.

5.2. Clinical Analysis

Based on the results of the data exploration we conducted an experiment that was more reflective of a clinical application of this approach. This required us to segment our data set into cross-validation groups based on individual babies. We then held out one baby's data as a test set and trained on the remaining 9 babies' data.

Using our previous techniques of support vector machines and decision trees we were able to obtain 90.2% and 89.8% accuracy respectively. However, accuracy is only

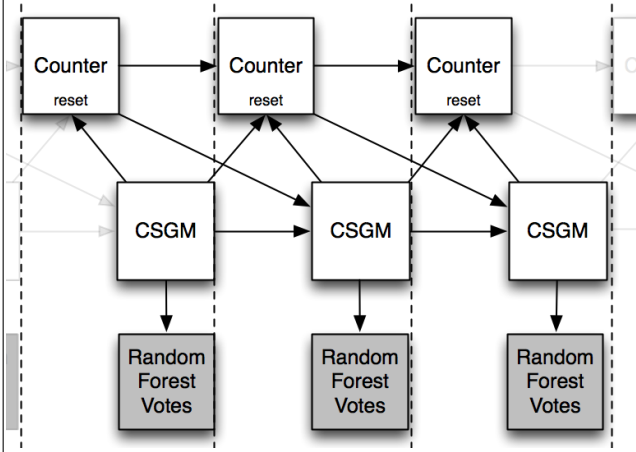


Figure 7. A model of cramped synchronized general movements using a Dynamic Bayes Net and a Random Forest.

one metric that needs to be analyzed in a clinical study, and despite these high numbers these techniques are not yet suitable for clinical use. The number of false positives and false negatives was too high for an effective medical device (see Table 1). Additionally evaluating this problem is made more complex by the requirement of recognizing events of distinct duration in a continuous stream of data samples. (Ward undertook a more thorough treatment of metrics [28])

Evaluating the results of these techniques showed that there was a tendency for the machine learning algorithm to change predictions of CSGMs more rapidly than was the case in the data. In order to attempt to compensate for this effect we also modelled this data using a Dynamic Bayes Net using the votes from a Forest of Random Decision Trees as observations [8].

Figure 7, shows the graphical model representation of the DBN that we used. The hidden binary variable, “CSGM” represents the presence or absence of a CSGM at that moment in time. There is one time-slice of the DBN for each accelerometer sample that was recorded. In order to reduce the dimensionality of our observations, we treated the number of votes from a Random Forest (using the full feature vector as input) as our observed variable. If all five trees in the forest, predicted the absence of a CSGM based on the calculated features, this node was a “0”. If all of the trees predicted the presence of a CSGM then the variable was a “5”. Partial votes resulted in integer values between 1 and 4.

We modeled the duration of a CSGM with a counter variable that incremented at each time step. If two successive time steps had a change in CSGM state, then the counter was reset to 1. The CSGM variable itself was updated based on the belief of the CSGM in the previous time slice and the

SVM		Classified as	
		Normal	Abnormal
Actual	Normal	598,914 (TN)	21,892 (FP)
	Abnormal	43,859 (FN)	3,291 (TP)

Decision Tree		Classified as	
		Normal	Abnormal
Actual	Normal	601,049 (TN)	38,883 (FP)
	Abnormal	25,126(FN)	2,898 (TP)

DBN + Random Forest		Classified as	
		Normal	Abnormal
Actual	Normal	451,039 (TN)	188,893 (FP)
	Abnormal	14,047 (FN)	13,977 (TP)

Table 1. Classification results from support vector machines, decision trees and dynamic bayes nets using random forests, broken down by true and false positives and true and false negatives (TP,FP,TN,FN).

length of time that the system had been in the normal or abnormal state. Based on the state of the CSGM a probabilistic distribution is then induced over the expected number of votes from the random forest.

This approach resulted in very different distribution of positive and negative predictions (see Table 1). A reduction in overall accuracy of about 20% was accompanied by more true positives, and a lower false positive rate. Depending on the relative cost of false positives and false negatives, different solutions may warrant different approaches. These trade-offs are shown visually in Figure 8. In this graph ideally a technique would plot in the upper left of the graph. The three techniques applied in the clinical analysis are shown along with straw-man techniques of always predicting the presence or absence of a CSGM.

6. Discussion and Conclusions

In this paper we have presented a method for recognizing cramped-synchronized general movements in babies who are born pre-term. CSGMs have been correlated with Cerebral Palsy when identified by a trained specialist from video. We have shown that a gesture recognition system has sufficient accuracy without requiring an hour of manual video annotation, to warrant further refinements in preparation for a clinical device.

Understanding how the machine learning techniques were able to achieve their accuracies is surprisingly hard and a focus of this paper. It is clear from our evaluation that there is a valuable signal that can be derived from combinations of features that is not present when features are evaluated independently. The best feature, the minimum over 2

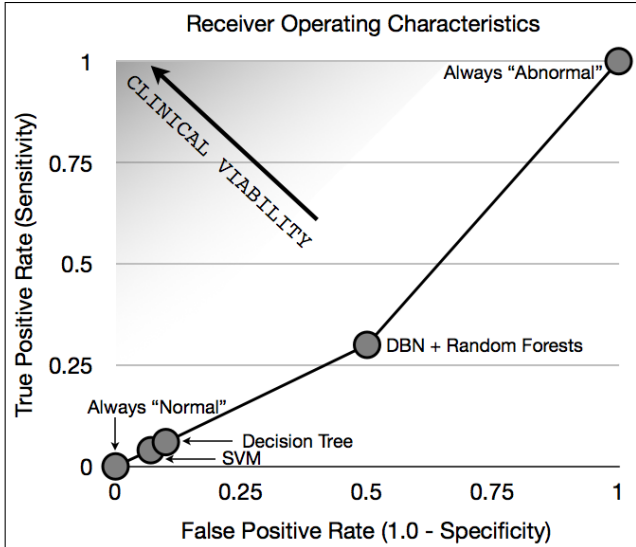


Figure 8. Receiver Operating Characteristics for the models tested.

seconds of the maximum of each sample, achieved significant gains over the majority vote, but the next six features also continued to provide additional improvements. Further metrics that explicitly characterize the reduction in work a video reviewer would need to do should be evaluated and are more clinically relevant.

The high correspondence between the trained nurse annotator and the machine learning technique supports existing inter-rater reliability studies [27]. Although somewhat inscrutable, the models developed in this study provide a quantitative translation of a gesture that can be recognized qualitatively and visually by people.

To improve the modeling of this data in the future, more informed distributions over the duration of a CSGM should be used. The current model implicitly enforces an exponential distribution over event lengths that is not optimal. Additionally the iterations of the DBN should be based on time, rather than samples (as some samples are missing). Further improvements may be gained by using the most informative features as observations instead of the votes from a random forest of decision trees.

There are clearly many more applications of subtle motion recognition that might have medical value if automatically detected. In relation to this study the “fidgety motion” that was also described in the literature as being anti-correlated with CP could also prove to have predictive power. It may be an important additional signal if larger clinical trials of this technique show greater variation among babies than was present in our cohort.

CP, like autism, is suspected to be an amalgamation of

several different underlying causes. Involuntary gesture recognition may provide a means by which different genres of diseases can be separated. Different types of CSGMs might indicate damage to different parts of the brain, or completely different types of injuries with similar manifestations. Such disease complexity argues for the use of multiple sensors even when they appear redundant, as the data may become less correlated for sub-types of a disease. A right arm sensor for example, might be able to pick out damage to a particular location of the brain, that a left arm sensor would miss.

Pentland in his book “Honest Signals” [23] refers to many quantifiably captured signals that humans involuntarily emit. Some of them reflect our emotions or role in a social network, but are unconscious or uncontrollable. Performing a similar evaluation as that done in this paper, of nuanced motions that people recognize visually, can’t stop themselves from performing, but aren’t well understood from a physical basis may help to illuminate their biological origins.

In this paper we successfully automated the identification of detailed kinetic physiological signals in newborn infants using wireless, lightweight accelerometers eliminating the need for a video observer and opening the opportunity for a new clinical intervention.

References

- [1] F. Albinali, M. S. Goodwin, and S. S. Intille. Recognizing stereotypical motor movements in the laboratory and classroom: a case study with children on the autism spectrum. In S. Helal, H. Gellersen, and S. Consolvo, editors, *UBICOMP*, pages 71–80, New York, NY, USA, 2009. ACM.
- [2] M. C. Allen and G. R. Alexander. Using gross motor milestones to identify very preterm infants at risk for cerebral palsy. *Developmental Medicine & Child Neurology*, 34(3):226–232, 1992.
- [3] O. Amft and G. Tröster. Recognition of dietary activity events using on-body sensors. *Artif. Intell. Med.*, 42(2):121–136, 2008.
- [4] R. Amstutz, O. Amft, B. French, A. Smailagic, D. Siewiorek, and G. Troster. Performance analysis of an hmm-based gesture recognition using a wristwatch device. In *CSE*, volume 2, pages 303–309. IEEE Computer Society, 29–31 2009.
- [5] D. Ashbrook and T. Starner. MAGIC: a motion gesture design tool. In E. D. Mynatt, D. Schoner, G. Fitzpatrick, S. E. Hudson, K. Edwards, and T. Rodden, editors, *CHI*, pages 2159–2168, New York, NY, USA, 2010. ACM.
- [6] M. Bachlin, D. Roggen, G. Troster, M. Plotnik, N. Inbar, I. Meidan, T. Herman, M. Brozgol, E. Shaviv, N. Giladi, and J. Hausdorff. Potentials of enhanced context awareness in wearable assistants for Parkinson’s disease patients with the freezing of gait syndrome. In *ISWC*, pages 123–130, 4–7 2009.

- [7] H. Brashear, T. Starner, P. Lukowicz, and H. Junker. Using multiple sensors for mobile sign language recognition. In S. Feiner and D. Mizell, editors, *ISWC*, page 45, Los Alamitos, CA, USA, 2003. IEEE Computer Society.
- [8] L. Breiman. Random Forests. *Machine Learning*, 45(1):5–32, 2001.
- [9] J. Cheng, O. Amft, and P. Lukowicz. Active capacitive sensing: Exploring a new wearable sensing modality for activity recognition. *Pervasive Computing*, pages 319–336, 2010.
- [10] L. E. Dunne, S. Brady, R. Tynan, K. Lau, B. Smyth, D. Diamond, and G. M. P. O’Hare. Garment-based body sensing using foam sensors. In *AUIC ’06: Proceedings of the 7th Australasian User Interface Conference*, pages 165–171, Darlinghurst, Australia, Australia, 2006. Australian Computer Society, Inc.
- [11] F. Ferrari, G. Cioni, C. Einspieler, M. F. Roversi, A. F. Bos, P. B. Paolicelli, A. Ranzì, and H. F. R. Prechtl. Cramped synchronized general movements in preterm infants as an early marker for Cerebral Palsy. *Arch Pediatr Adolesc Med*, 156(5):460–467, 2002.
- [12] R. L. Goldenberg, J. F. Culhane, J. D. Iams, and R. Romero. Epidemiology and causes of preterm birth. *The Lancet*, 371(9606):75 – 84, 2008.
- [13] H. Junker, O. Amft, P. Lukowicz, and G. Trster. Gesture spotting with body-worn inertial sensors to detect user activities. *Pattern Recognition*, 41(6):2010 – 2024, 2008.
- [14] K. Lyons, H. Brashear, T. Westeyn, J. Kim, and T. Starner. GART: The gesture and activity recognition toolkit. *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*, pages 718–727, 2007.
- [15] C.-H. Min, Y. Kim, A. Tewfik, and A. Kelly. Detection of self-stimulatory behaviors of children with autism using wearable and environmental sensors. *Journal of Medical Devices*, 3(2):027506, 2009.
- [16] D. Minnen, T. Starner, M. Essa, and C. Isbell. Discovering characteristic actions from on-body sensor data. In *ISWC*, pages 11–18. IEEE, 11-14 2006.
- [17] P. Mistry, P. Maes, and L. Chang. WUW - wear ur world: a wearable gestural interface. In D. R. O. Jr., R. B. Arthur, K. Hinckley, M. R. Morris, S. E. Hudson, and S. Greenberg, editors, *CHI Extended Abstracts*, pages 4111–4116, New York, NY, USA, 2009. ACM.
- [18] K. B. Nelson and J. H. Ellenberg. Neonatal signs as predictors of Cerebral Palsy. *Pediatrics*, 64(2):225–232, 1979.
- [19] N. I. of Health. National Institute of Neurological Disorders and Stroke, May 2010.
- [20] C. Park and P. Chou. Eco: Ultra-wearable and expandable wireless sensor platform. In *BSN*, pages 165–168. IEEE Computer Society, 3-5 2006.
- [21] S. Patel, K. Lorincz, R. Hughes, N. Huggins, J. Growdon, D. Standaert, M. Akay, J. Dy, M. Welsh, and P. Bonato. Monitoring motor fluctuations in patients with Parkinson’s disease using wearable sensors. *Trans. Info. Tech. Biomed.*, 13(6):864–873, 2009.
- [22] D. J. Patterson, D. Fox, H. A. Kautz, and M. Philipose. Fine-grained activity recognition by aggregating abstract object usage. In K. Mase and B. Rhodes, editors, *ISWC*, pages 44–51, Osaka, Japan, October 2005. IEEE Computer Society.
- [23] A. S. Pentland. *Honest Signals: How They Shape Our World*. The MIT Press, 2008.
- [24] S. Saigal and L. W. Doyle. An overview of mortality and sequelae of preterm birth from infancy to adulthood. *The Lancet*, 371(9608):261 – 269, 2008.
- [25] T. S. Saponas, D. S. Tan, D. Morris, R. Balakrishnan, J. Turner, and J. A. Landay. Enabling always-available input with muscle-computer interfaces. In A. D. Wilson and F. Guimbretière, editors, *UIST*, pages 167–176, New York, NY, USA, 2009. ACM.
- [26] E. Sazonov, N. Sazonova, S. Schuckers, M. Neuman, and C. S. Group. Activity-based sleep–wake identification in infants. *Physiological Measurement*, 25(5):1291, 2004.
- [27] L. M. Snider, A. Majnemer, B. Mazer, S. Campbell, and A. F. Bos. A comparison of the general movements assessment with traditional approaches to newborn and infant assessment: Concurrent validity. *Early Human Development*, 84(5):297 – 303, 2008.
- [28] J. Ward, P. Lukowicz, and G. Tröster. Evaluating performance in continuous context recognition using event-driven error characterisation. In M. Hazas, J. Krumm, and T. Strang, editors, *LoCA*, volume 3987 of *Lecture Notes in Computer Science*, pages 239–255. Springer, 2006.
- [29] A. Zinnen, K. van Laerhoven, and B. Schiele. Toward recognition of short and non-repetitive activities from wearable sensors. *Ambient Intelligence*, pages 142–158, 2007.