Effects of Intra-Subject Variation in Gait Analysis on ASD Classification Performance in Machine Learning Models

Benn Henderson
School of Computing,
Engineering and Intelligent
Systems
Ulster University
Londonderry, Northern
Ireland
henderson-b7@ulster.ac.uk

Pratheepan Yogarajah
School of Computing,
Engineering and Intelligent
Systems
Ulster University
Londonderry, Northern
Ireland
p.yogarajah@ulster.ac.uk

Bryan Gardiner
School of Computing,
Engineering and Intelligent
Systems
Ulster University
Londonderry, Northern
Ireland
b.gardiner@ulster.ac.uk

Martin McGinnity
School of Computing,
Engineering and Intelligent
Systems
Ulster University
Londonderry, Northern
Ireland
tm.mcginnity@ulster.ac.uk

Kitty Forster
School of Psychology
Bangor University
Bangor, UK
k.forster@bangor.ac.uk

Bradley Nicholas
School of Psychology
Bangor University
Bangor, UK
b.nicholas@bangor.ac.uk

Dawn Wimpory
School of Psychology
Bangor University
Bangor, UK
d.wimpory@bangor.ac.uk

Jithangi Wanigasinghe
Department of Paediatrics
University of Colombo
Sri Lanka
jithangi@pdt.cmb.ac.lk

Abstract—Autism Spectrum Disorder (ASD) is a developmental disorder that is prevalent globally. Research into detecting autism traditionally focused on behavioural aspects of the condition, however, more recently, focus has shifted to more objective alternatives using techniques such as machine learning and gait analysis. Gait measurements, having been used for person identification, varies from person to person, introducing a lot of intra-subject variance. This applies to the 8 spatial-temporal features used in this study, representing the time that an individual spends in each phase of a gait cycle, collected using a Vicon motion tracking system. The features were averaged across each gait trial that the subjects performed, producing a second set of features with reduced intra-subject variance. Four common classifiers, a Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Random Forests (RF) and a Decision Tree (DT) classifier, were all trained using the two feature sets and their classification rates were compared. The results show that for the RF classifier, reducing the intra-subject variance, was able to successfully increase the classification power. The KNN and DT classifiers experienced a minimal decrease in accuracy, where the SVM suffered the greatest loss when intra-subject variance was reduced. Results overall show that the effect intra-subject variance has on classification power depends heavily on the suitability of the classifier to the initial problem as well as size and class balance of the data.

Keywords—Machine Learning, Autism Spectrum Disorder, automatic classification, intra-subject, variation.

I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a permanent neurological condition that is prevalent in the entire world population [1]. Despite the field focusing on the behavioural aspect of ASD [2], [3], more recent research has been conducted indicating that ASD also influences the motor control for many people with the condition. This can manifest itself through various, measurable features relevant to a person's gait such as the kinematics of their joints [4] and ground reaction forces during perambulation [5]. What makes these measurements powerful for identifying ASD is that they are objective measurements that can be recorded using accurate tracking analysis tools [6], [7]. Objective measurements have the potential to identify ASD with more confidence when compared to behavioural assessments recorded by human beings.

To date, gait analysis has been utilised extensively for security applications, for example, person detection and tracking [8] or as a form of person identification [9] [10]. Implications of this suggest that gait representations are unique to individual people and can be used to extract identifiable biometric information from someone's recorded gait. A number of works using gait analysis utilise motion tracking systems such as the Vicon infra-red camera system, where markers placed on a person's lower limbs can measure useful information about their gait or how they walk [7], [11]. Positional information can be used to determine which phase of the gait cycle, defined as the time period in which one foot contacts the ground to when that same foot contacts the ground again, someone is in as well as detecting gait events such as when the foot strikes the ground [6].

Studies have investigated different combinations of features related to gait in the identification of autism and other conditions. In [6], the Vicon and force plates were used to gather vertical ground reaction forces as well as temporal features such as stance time. These features were tested for significance using t-tests and Mann Whitney-U tests, looking for differences in those with autism and those without. In the mentioned study, each subject was represented by a single gait cycle.

The utilisation of machine learning paradigms for automating classification problems has become common practice [12], [13]. Support Vector Machines (SVM) and K-Nearest Neighbours (KNN) are both commonly applied models when dealing with binary classification problems [14], [15], and for this reason, they act as good baseline classifiers for comparison and analysis of newly proposed models. Furthermore, they are often well supported in machine learning programming libraries, providing highly efficient model implementations requiring little demand on processing power, making them popular among researchers and developers. For autism classification, other classifiers such as Random Forests (RF) and Decision Trees (DT) have also proven to be effective in achieving high classification results, with [13] making use of RF, KNN and SVM classifiers on adolescent scan data to detect autism, achieving 95%, 89% and 100% respectively. In [16], random regression forests, c4.5 decision trees and PART classifiers were used on gaze and demographic feature data to detect autism

resulting in the PART classifier having the best accuracy at 96.2%. In [17], spatial-temporal gait features were applied to RFs, SVMs and a Kernel Fisher Discriminant (KFD), with the purpose to perform classification to detect those with Parkinson's Disease from those without. In [18], force plates were used to gather spatial-temporal, kinetic and kinematic gait parameters before applying them to an SVM and an Artificial Neural Network (ANN) to perform automatic autism classification.

Although existing literature provides promising results for automatic classification of autism and related conditions, their focus is often on the types of features [6], classification of other conditions such as Parkinson's Disease [17] or classification comparison between machine learning techniques [18]. The present study expands on these studies by looking at how each subject is represented in the data and how this impacts classification.

Considering that gait can be used to differentiate some people with ASD from those without, there still must exist a lot of variance within these groups. This is potentially amplified by how a person's gait itself can vary for each gait cycle. This variance is noteworthy when reviewing work conducting autism classification, datasets collected tend to use a small number of subjects. In both [5] and [6] for example, a dataset containing 60 subjects overall was used. 30 of these subjects had ASD and 30 were typically developed. Although this indicates a good balance between the groups, due to large variation in gait, this dataset will only represent a small portion of people, both with and without autism. It is often more common to have less balanced datasets such as in [18], where 12 of the subjects have ASD and 32 were typically developed. To be able to effectively distinguish between subjects and therefore groups of people with ASD and without, the goal would be to reduce intra-subject variance.

One way to achieve a representation that reduces the variance between gait cycles, is to average the features obtained during walking trials, that represent individual gait cycles, across a single trial instead. The present study is therefore designed to check how this change in representation can affect classification rates of autism when using machine learning algorithms. Results from such a study could provide useful insight into how to further explore the data used for these types of studies.

II. METHODOLOGY

A. Data Acquisition

Using the Vicon motion tracking system, gait data were collected from 21 subjects, 12 of these subjects clinically diagnosed as having ASD, while the remaining 9 were typically developed (TD). The raw data consisted of the number of frames from the beginning of the walking trial, sampled at a frequency of 250Hz. Every subject had completed a minimum of five walking trials each, with each walking trial consisting of a varying number of gait cycles. Within each gait cycle, the frame number between each gait even was recorded, and the difference in frame number between two gait events, making a gait phase, was used as the initial feature set.

Two sets of features were generated from the gait phase timings. The first (Dataset A), consisted of limiting each subject's data to only include the five minimum common number of trials, which resulted in 167 gait cycles representing subjects with ASD, and 111 gait cycles representing the TD subjects. The second set of features (Dataset B), comprised of the average of each gait phase timing across all gait cycles for each trial to produce a single data point representing that trial, again using the five minimum common number of trials per subject, which resulted in 60 gait cycles representing the subjects with ASD and 45 gait cycles representing the TD subjects.

B. Data Pre-processing

Within each trial, the gait cycle includes valuable temporal information and eight features are generated to represent a gait cycle. The eight features being used are temporal features that correlate to the time spent in different phases of the gait cycle. The 4 phases are; Step Time, Stance Time, Swing Time and Step to Step. For each phase, two features are extracted as the phases repeat for each leg. If 'Right', 'Left', 'Strike' and 'Off' were denoted as 'R', 'L', 'S' and 'O' respectively, all 8 features can be represented. For example, Right Strike (RS) is when the right foot strikes the ground and Left Off (LO) is the time when the left foot lifts off of the ground. The full set of eight features, along with their corresponding phase, can be represented as follows; RS-LS, LS-RS (Step Time), LO-LS, RO-RS (Stance Time), LS-LO, RS-RO (Swing Time), LS-LS, RS-RS (Step to Step). Each value was calculated as the difference in frames from when the two events are recorded using the Vicon.

As different features represent different phases of the gait cycle, and some phases naturally last longer than others, the range of values that can be collected for these will differ. By normalizing the values, the collected data is converted into one of a common scale without changing relationships between them. Normalisation was conducted by scaling each feature vector individually to the unit norm. For this study, the 'l2' or 'Euclidean' norm was used.

C. Classification

Four machine learning algorithms were chosen based on their track record for performing well within the field of ASD classification. As commonly found within trials similar to this, the small dataset size, may often produce wildly different results between different algorithms. As the goal of the study is to determine the effect of intra-class variance on the classification performance in general, applying the process to four different classifiers will allow us to investigate how different classifiers react to the same small dataset and these temporal features. This will be useful in other studies investigating gait and ASD classification. Each of the four machine learning algorithms will be processed and trained following the same process described in the remainder of this section.

The four machine learning models being used are K-Nearest Neighbours (KNN), Support Vector Machines (SVM), Random Forests (RF) and Decision Trees (DT). KNN classifiers are conceptually simple to understand and are easy to implement, and are being used here as a baseline for comparison. As the study is using a small dataset, the classifier shouldn't suffer from long computational times that are associated with larger datasets when calculating the distance between all the data points. SVMs are a strong classifier, often used in small datasets. They have also been commonly applied to ASD classification problems. Considering this, results from this study could potentially support their default use in the

field. RFs and DTs are less common when dealing with autism classification; however, they have been used in several studies, producing motivating results. As both RFs and DTs require branching when generating classification decisions, they can often perform well with problems where the data isn't as linearly separable. Using these four classifiers, will represent a considerably large set of machine learning techniques and this will, hopefully, translate to a thorough and useful comparison of the effects of intra-subject variance in small datasets.

D. Cross Validation

In order to validate the results, such that the performance can be estimated when given unseen data, cross validation is applied to all models that are being trained. 10-fold cross validation is chosen as, although the dataset is small enough so that leave-one-out cross validation is plausible, there will still exist some relation between the data in the training and testing splits. This is because, as each gait cycle in the first set of tests is being used as an individual data point to be classified, there is a situation where the training data and the testing data can both contain data points from the same subject and even from the same trial. The impact this will have on the results is potentially limited, as there will still exist variance in the features for each gait cycle and between each trial. Therefore, cross-validation will still estimate how effective the classifiers will perform with new data points. When using Dataset B, where the features are averaged for each trial, and the trial is then used as a single data point, the impact will be reduced even more. This is because, the variance that exists between each gait cycle is now captured among one trial and one data point compared to the first set of tests. This reduction in intra-class variance should allow classifiers to identify the difference between the classes much easier and should therefore improve the prediction power of the classifiers.

E. Grid Search

Each of the four classifiers will have several parameters that can be adjusted depending on the problem at hand, having a massive impact on the final performance of the produced classifier. This presents a problem, in that if optimal parameters are chosen for one classifier yet sub-optimal parameters are chosen for another, then the results will likely always favour the classifier with the optimal parameters. This is under the assumption that all four classifiers are similar in their natural fitness for the problem. To address this, multiple values for each parameter for each classifier will be chosen using a common range of values for that parameter, and a grid search will be conducted for each classifier. Cross-validation, as explained in the previous section, will be applied for each combination of parameters so that when the models are compared, the comparison is between the estimated performances when considering new data. Ultimately, a new model is produced for each combination of parameters to determine the optimal model for that classifier, which can then be compared with the optimal model produced for each of the other classifiers. As part of the grid search implementation, a metric must be supplied to decide which model is considered a better classifier. This can change, depending on the application, however in this case, when the application is likely to be used to simply identifying people with ASD from those without, precision has been chosen. Although precision is used here, the analysis section will cover all four metrics; namely, accuracy, f1 score, recall and precision.

III. RESULTS

A. Data Variance

The purpose of this study is to investigate the effect reducing the intra-subject variance has on classification power. To achieve a better understanding of how the intra-class variance is affected, some of the data was used to generate scatter graphs that represent a subset of the features for a subset of the subjects, before and after the mean average feature values were used to represent each trial.

Considering Dataset A, Fig. 1 plots two of the eight features from this dataset against each other, with each point representing a single gait cycle. The two features chosen for this were the first two features that appear in the data set, RS-LS and LS-RS. Fig. 2 plots the averaged versions of the same two features (Dataset B), across the same value range for comparison. Here each point represents a full gait trial.
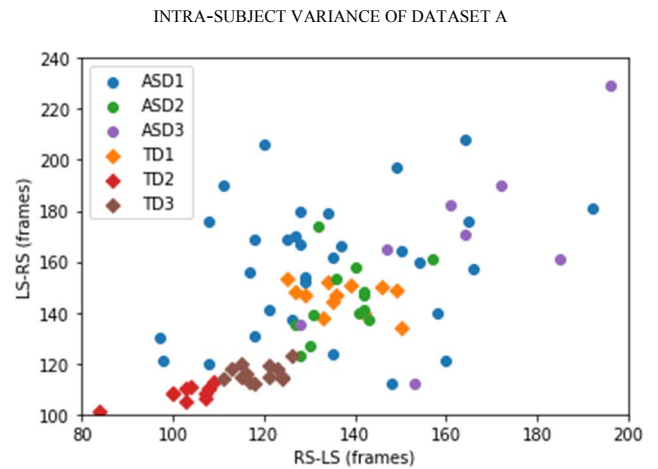
INTRA-SUBJECT VARIANCE OF DATASET A

Fig. 1. Scatter graph showing the variance in LS-RS and RS-LS from Dataset A between gait cycles for 3 subjects with ASD and 3 subjects that are TD.
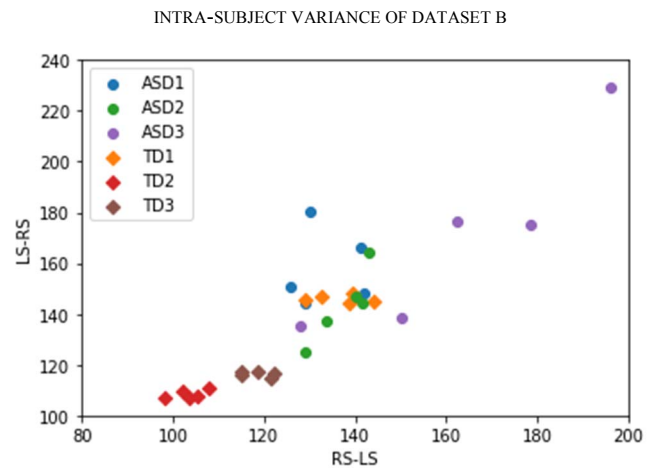
INTRA-SUBJECT VARIANCE OF DATASET B

Fig. 2. Scatter graph showing the variance in LS-RS and RS-LS from Dataset B between gait trials for 3 subjects with ASD and 3 subjects that are TD.

For both figures, the same six subject's data were used and are represented in the same way. Three of the subject's data that was visualised were subjects with ASD, and the other three were TD subjects. ASD subject's data are represented using the circles, and the TD subject's data are

represented using the diamonds as indicated through the legend.

Two main characteristics of these graphs that should be noted, and discussed further below, are; the amount of data being used to represent each subject, and the variance within each subject's data as displayed.

The difference in the number of points displayed in Fig. 1 compared to the number of points displayed in Fig. 2, with there being much more in Fig. 1, than in Fig. 2, is indicative of how averaging has reduced the amount of data that has to be processed by a machine learning algorithm. Considering the situation where a model performs better using data represented in Fig. 2 than it does using data in Fig. 1, then the classification power has not only been increased, but it will have been achieved using much less data.

For each subject's data in Fig. 1, it is clearly visible that the range of values covered by the points is much larger than in Fig. 2. For 'ASD1', the difference in variance in its representation is the most extreme. From covering nearly, the entire span of both axes, to covering only a small portion. This confirms that the averaging technique, on some level, has reduced intra-subject variance, in some cases, to a significant level. For other subjects such as 'TD2' and 'TD3', the difference in variance seems much smaller, although in these cases, their variance was also smaller than the others in Fig. 1 before any averaging had taken place. Therefore, the averaging technique has been shown to reduce intra-subject variance more so in data that contained more variance and had less impact on data whose variance was already limited.

In addition, a key feature that seems evident in both Fig. 1 and Fig. 2 is that the amount of variance for subjects with ASD is much greater than in the subjects that are TD. This could imply that those with ASD, naturally vary their gait much more than those who are typically developed. It could potentially prove beneficial then, to include objective measurements of the variance of key gait features, in the feature vectors themselves as a method of improving classification power. This is, however, out of the scope of the present study and can be explored in future work.

B. Classification Results

In this section, results are presented separately for Dataset A, with each gait cycle as a feature vector, and Dataset B, averaging each feature for each gait cycle contained within the trial as a feature vector. Four classifiers were trained for each dataset so that resultant metrics can be compared for different commonly used classifiers. These tests were chosen so that the impact that reducing the intra-subject variance has on classification power can be investigated. Accuracy, recall, precision and F1 Score (weighted average of Precision and Recall) were calculated; descriptions of each metric found by reviewing the 'metrics' module documentation from the 'sklearn' python library. Accuracy and F1 Score are being used as the main criteria for determining the classification power, making use of recall and precision for extended analysis.

All eight temporal features are used to train each classifier. Cross validation and grid search were used to produce the resulting metrics, thus validating the performance of the classifiers on unseen data while the grid search produces the classifier with the optimal parameters for the problem.

Table. I shows the metrics for each classifier that was trained using the features from each individual gait cycle and Table II shows the metrics for the classifiers that were trained on the averaged features for each trial instead. The metrics are presented as percentage values between 0 and 100.

TABLE I. CLASSIFICATION RESULTS USING DATASET A

Classifier	Features			
	Accuracy (%)	Recall (%)	Precision (%)	F1 Score (%)
SVM	63.75	69.57	85.71	76.80
KNN	58.75	78.05	57.14	65.98
RF	68.75	80.39	73.21	76.63
DT	63.75	76.47	69.64	72.90

TABLE II. . CLASSIFICATION RESULTS USING DATASET B

Classifier	Features			
	Accuracy (%)	Recall (%)	Precision (%)	F1 Score (%)
SVM	50.00	100.00	50.00	66.67
KNN	56.67	40.00	60.00	48.00
RF	76.67	86.67	72.22	78.79
DT	63.33	60.00	64.29	62.07

Fig. 3 displays the accuracy and F1 Score metrics for each of the classifiers trained on Dataset A in a bar graph, allowing us to more easily compare them. Fig. 3 and Fig. 4 don't display recall and precision metrics as the F1 Score captures the key characteristics of both.

From Fig. 3, the RF classifier has been chosen as the preferred classifier, closely followed by the SVM. The reason the RF was chosen over the SVM is due to the higher accuracy, despite having similar F1 Scores. The RF reported an accuracy of 68.75%, a recall of 80.39% and a precision of 73.21%, compared to the SVM which had an accuracy of 63.75%, a recall of 69.57% and a precision of 85.71%.

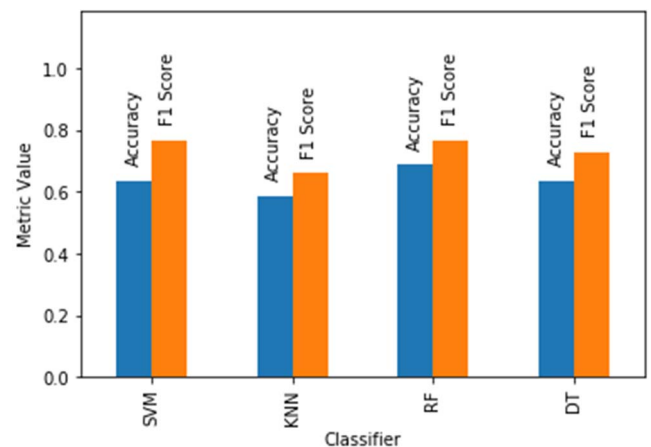CLASSIFICATION POWER USING DATASET A

Fig. 3. Bar graph representation of the results from training the 4 classifiers on features from each gait cycle.

Fig. 4, as in Fig. 3, displays the accuracy and F1 Score metrics from the best of each classifier, this time using Dataset B. This graph shows that the RF classifier again, has

the best classification power. With an accuracy of 76.67%, recall of 86.67% and a precision of 72.22% the RF completely outperforms the other classifiers. With an F1 score of 78.79%, the next best classifier for Dataset B was the SVM with an F1 Score of 66.67%. For accuracy, the next best classifier was the DT classifier at 63.33%.
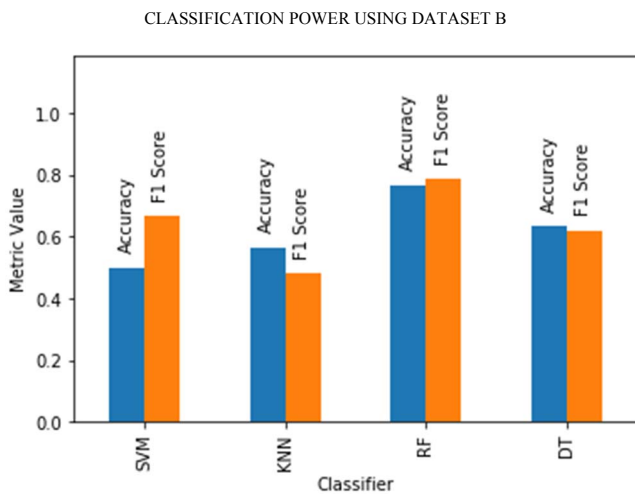
CLASSIFICATION POWER USING DATASET B

Fig. 4. Bar graph representation of the results from training the 4 classifiers on features from averaging gait cycles per trial.

C. Summary of Key Points

The process of reducing the intra-subject variation had a positive impact on the RF classifier with an increase in both the accuracy and F1 Score. Although the increase in F1 Score was not large (76.64% - 78.79%), the increase in accuracy was a much more significant positive change (68.75% - 76.67%).

The KNN and DT classifiers both observed minimal changes in their accuracy between being applied to Dataset A and Dataset B. Their F1 Scores dropped significantly however, indicating a decrease in classification power and a change in which subjects are being correctly and incorrectly classified.

The SVM classifier also experienced a negative impact due to the averaging of the gait trials. The model went from providing the best F1 Score and a promising accuracy when applied to Dataset A to producing a meaningless classifier that always predicts the data to be from an autistic subject when using the averaged trial versions of the features (Dataset B). This is possibly explained by the difference in the amount of data between the two tests, as by averaging multiple gait cycles in a trial to produce a single value, the overall information used in training is being limited.

D. Positive Impact Case

The improvement in the RF classifier between the two tests provides a case for the positive impact that reducing the intra-subject variation in gait data has on ASD classification power. Despite the large increase in accuracy, the increase observed for the F1 Score was much smaller.

The smaller increase for the F1 Score can be explained through the change in individual recall and precision metrics between tests. The reduction in precision from 73.21% to 72.22% was traded for the much larger increase in recall from 80.39% to 86.67%.

The difference in magnitude of changes between these two metrics, could be an indicator that there are external factors also impacting the classification power of the models. One possible factor could be the imbalance between the two classes being represented in the collected data. This may be due to the use of the averaging technique also reducing the amount of data being used to train the classifiers, which may be highlighting the impact of the imbalance between the two classes.

E. Minimum Impact Cases

The KNN and DT classifiers observed a minimal negative trend in their accuracy metrics. For the KNN classifier this change was less than 3% and for the DT the change was less than 1%. This minimal change in accuracy implies that the reduction in intra-subject variance had limited impact on classification power.

In both tests, the KNN and DT classifiers results were weaker than that of the SVM when applied to Dataset A and that of the RF when applied to both Dataset A and B. Through using the grid search, an optimal version of each classifier is being compared. Knowing this, it can be confidently said that the KNN and DT classifiers are not suited to this problem using this data set and that the negative effect the reduction in intra-class variation had, may be more indicative of external factors influencing the results.

F. Outlying Case

When applying Dataset A, it is noted that the SVM classifier had a slightly higher F1 Score than the RF classifier in the same test. With the difference being miniscule, and the SVM having a lower accuracy score, the conclusion is that the RF classifier is a better choice under the conditions of the data used for this study.

The results from applying Dataset B show that the performance of the SVM classifier was decreased significantly by the reduction in intra-subject variation. The magnitude of the change is highlighted when looking at both the recall and precision scores. The recall had risen to 100% while the precision fell to 50%. Such an extreme reaction to the change in the data, prompted further investigation leading to the revelation that the classifier was predicting that every data point was of the ASD class. A classifier predicting only one class has no use in real world applications.

G. Conclusion of Results

The averaging technique used to reduce intra-subject variance proved to have a positive impact on the RF classifier, increasing both its accuracy and F1 Score. Despite this positive outcome, the other classifiers such as the KNN and DT classifiers, had minimal changes to accuracy and experienced a more negative impact on the F1 Score. However, both classifiers produced weaker results when used with Dataset A, so there is a case to be made about external factors dominating the effect on performance for these classifiers. Similarly, for the SVM, although providing a classifier in contention for being the best of all four classifiers in the first test, after applying the averaging technique, the SVM classifier produced was only predicting ASD for all new data. It will therefore be important to consider the imbalance in the data or the use of small data sets when deciding if the averaging technique will be beneficial for a classification problem.

IV. Discussion

The purpose of this research was to compare the effect of different intra-subject variance in gait when applied to autism classification using machine learning. Data was collected using a Vicon motion tracking system, resulting in temporal measurements of gait events called here 'strike' and 'off' for each foot. Indicating the point in time where one of the feet struck the ground after being in the air, and where one of the feet was initially lifted off the ground. These measurements were converted to temporal features representing the time spent in different gait phases for each gait cycle recorded.

During training of the classifiers, a combination of cross-validation and a grid search were used to find the optimal parameters for each classifier. Classification itself was then completed using the best performing model from each tested algorithm. Results were compared using 4 metrics, accuracy, F1 Score, recall and precision.

Datasets found in the existing literature mentioned thus far, contain varying quantities of subjects and differing levels of imbalance between the classes. Comparing these datasets to the dataset being used in the current study, with 12 ASD subjects and 9 typically developed, the impact of variation in the subjects may be magnified due to the large portion of the dataset each subject represents. The results from this study can, therefore, provide further context when handling smaller autism related, gait analysis datasets.

Further investigation into the SVM classifier trained using Dataset B, found that it was predicting ASD every time, possibly due to class imbalance. Analysing the difference between Dataset A and Dataset B, it can be said that while reducing the amount of variation in data representing a single subject, there was also an increase in the imbalance of the data overall. Other methods to counter this can be explored, such as algorithms to produce artificial data for the class that has fewer data points to remove the imbalance in the smaller dataset.

V. Conclusion

The effect that the amount of variation has on prediction power was studied in this preliminary investigation. It was found that by reducing the intra-subject and intra-class variation, the classification power of the RF classifier can be greatly improved. The trade-off is that other classifiers, such as the KNN and DT classifiers may respond negatively to the variation and data set size change. Further investigations should look at the effects of the reduction in variation when used in combination with data generation techniques to remove imbalance in datasets and prevent overfitting as was witnessed with the SVM classifier. After reducing the variation by averaging the gait cycles, the portion of the full dataset that each data point represented increased, causing the SVM classifier to only predict ASD due to this being enough to get good results during training. In cases such as this, it was found that the reduction in variation is not worth the loss of the dataset size.

References

[1] M. Elsabbagh *et al.*, "Global Prevalence of Autism and Other Pervasive Developmental Disorders," *Autism Res.*, vol. 5, no. 3, pp. 160–179, 2012, doi: 10.1002/aur.239.

[2] J. Hashemi *et al.*, "A computer vision approach for the assessment of autism-related behavioral markers," *2012 IEEE Int. Conf. Dev. Learn. Epigenetic Robot. ICDL 2012*, pp. 1–7, 2012, doi: 10.1109/DevLrn.2012.6400865.

[3] B. Wingfield *et al.*, "A Predictive Model for Paediatric Autism Screening," *Health Informatics J.*, pp. 1–12, 2019, doi: 10.1177/ToBeAssigned.

[4] M. Calhoun, M. Longworth, and V. L. Chester, "Gait patterns in children with autism," *Clin. Biomech.*, vol. 26, no. 2, pp. 200–206, 2011, doi: 10.1016/j.clinbiomech.2010.09.013.

[5] C. Z. C. Hasan, R. Jailani, and N. M. Tahir, "ANN and SVM Classifiers in Identifying Autism Spectrum Disorder Gait Based on Three-Dimensional Ground Reaction Forces," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2018-Octob, no. October, pp. 2436–2440, 2019, doi: 10.1109/TENCON.2018.8650468.

[6] C. Z. C. Hasan, R. Jailani, N. M. Tahir, and H. M. Desaa, "Vertical ground reaction force gait patterns during walking in children with autism spectrum disorders," *Int. J. Eng. Trans. B Appl.*, vol. 31, no. 5, pp. 705–711, 2018, doi: 10.5829/ije.2018.31.05b.04.

[7] C. Z. C. Hasan, R. Jailani, and N. Md Tahir, "Use of statistical approaches and artificial neural networks to identify gait deviations in children with autism spectrum disorder," *Int. J. Biol. Biomed. Eng.*, vol. 11, pp. 74–79, 2017.

[8] L. Wang, H. Ning, T. Tan, and W. Hu, "Fusion of static and dynamic body biometrics for gait recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 2, pp. 149–158, 2004, doi: 10.1109/TCSVT.2003.821972.

[9] N. Kala, T. Bhatia, and N. Aggarwal, "Person Identification and Characterization from Gait Using Smartphone," *2019 11th Int. Conf. Commun. Syst. Networks, COMSNETS 2019*, vol. 2061, pp. 492–495, 2019, doi: 10.1109/COMSNETS.2019.8711131.

[10] M. Babaee, L. Li, and G. Rigoll, "Person identification from partial gait cycle using fully convolutional neural networks," *Neurocomputing*, vol. 338, pp. 116–125, 2019, doi: 10.1016/j.neucom.2019.01.091.

[11] P. Merriaux, Y. Dupuis, R. Boutteau, P. Vasseur, and X. Savatier, "A study of vicon system positioning performance," *Sensors (Switzerland)*, vol. 17, no. 7, pp. 1–18, 2017, doi: 10.3390/s17071591.

[12] S. Shetty and Y. S. Rao, "SVM based machine learning approach to identify Parkinson's disease using gait analysis," *Proc. Int. Conf. Inven. Comput. Technol. ICICT 2016*, vol. 2, pp. 1–5, 2017, doi: 10.1109/INVENTIVE.2016.7824836.

[13] A. Demirhan, "Performance of Machine Learning Methods in Determining the Autism Spectrum Disorder Cases," *Mugla J. Sci. Technol.*, pp. 79–84, 2018, doi: 10.22531/muglajsci.422546.

[14] T. T. Verlekar, L. D. Soares, and P. L. Correia, "Automatic classification of gait impairments using a markerless 2D video-based system," *Sensors (Switzerland)*, vol. 18, no. 9, pp. 1–16, 2018, doi: 10.3390/s18092743.

[15] C. Z. C. Hasan, R. Jailani, N. M. Tahir, and R. Sahak, "Autism spectrum disorders gait identification using ground reaction forces," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 15, no. 2, pp. 903–911, 2017, doi: 10.12928/TELKOMNIKA.v15i2.6143.

[16] S. Canavan *et al.*, "Combining gaze and demographic feature descriptors for autism classification," in *Proceedings - International Conference on Image Processing, ICIP*, 2018, doi: 10.1109/ICIP.2017.8296983.

[17] F. Wahid, R. K. Begg, C. J. Hass, S. Halgamuge, and D. C. Ackland, "Classification of Parkinson's disease gait using spatial-temporal gait features," *IEEE J. Biomed. Heal. Informatics*, vol. 19, no. 6, pp. 1794–1802, 2015, doi: 10.1109/JBHI.2015.2450232.

[18] S. Ilias, N. M. Tahir, R. Jailani, and C. Z. C. Hasan, "Classification of autism children gait patterns using Neural Network and Support Vector Machine," *ISCAIE 2016 - 2016 IEEE Symp. Comput. Appl. Ind. Electron.*, pp. 52–56, 2016, doi: 10.1109/ISCAIE.2016.7575036.