Using Acoustic Transmission Features and Decision Trees for Classification of Developmental Dysplasia of the Hip

M. Hossen[1], R. Sandler[1,2] and H. Mansy[1,2]

1. University of Central Florida, Orlando, Florida, USA
2. Biomedical Acoustic Research Company, Orlando, Florida, USA
mdrakib.hossen@ucf.edu, rhsandler@gmail.com, Hansen.Mansy@ucf.edu

Introduction: Developmental dysplasia of the hip (DDH) is a pediatric condition where the hip joint is improperly formed resulting in abnormalities in femoral head, acetabulum or both, which can lead to mild acetabular shallowness or complete dislocation [1]. DDH incidence may be as high as 4% to 6% in newborn infants[2], [3]. Prompt treatment such as Pavlik Harness or casting may ensure normal hip development and prevent long term disability. The older the age at DDH presentation, the worse the outcomes after intervention. As the infant grows, the non-invasive treatments often become ineffective necessitating surgical interventions, with generally poorer outcomes [4], [5].

Current newborn screening relies on physical examination (e.g., the Ortolani/Barlow tests), which are recommended to be performed within 3 months of birth, but accuracy depends on expert performance [1], [6]. In addition, mild acetabular dysplasia without instability may yield false negative results on these examinations [4]. The Ortolani and Barlow maneuvers have low sensitivity (~36%) but high specificity (~98%) [12]. Further diagnosis using ultrasonography (US) may be required — potentially on a periodic basis, if other abnormal physical findings or risk factors (e.g., breech presentation, positive family history, female sex, or firstborn status) show potential sign of DDH.

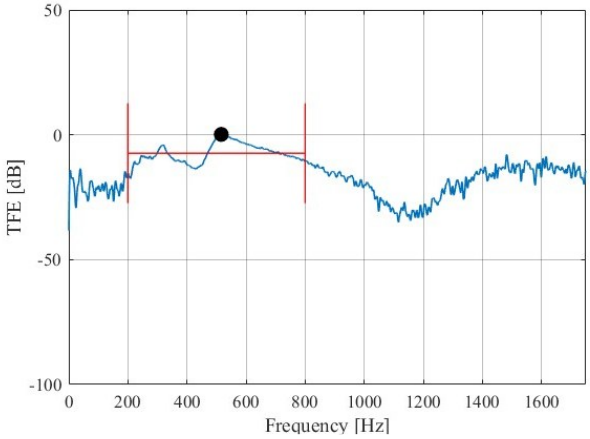

Figure 1. Left-to-right transfer function (TFE) of a normal subject. The horizontal red line represents the average value in 200-800 frequency range. Black dot is the maximum value within the range. The crest (aka. 'mx') is calculated as the maximum – average, in the selected frequency range.

Ultrasound screening in one study showed sensitivity of 88.5 % and specificity of 96.7 % [13]. But US may not be accessible in all healthcare environments, especially for screening in primary care environments. Cost and lack of highly skilled professionals also hinders US availability on regular basis, especially in low-technology settings. An alternative non-invasive, inexpensive, and easy-to-use methods that require minimal expertise could significantly reduce healthcare burden and improve outcome.

One promising approach is the use of acoustic transmission to detect structural abnormalities in the hip joint. Several studies have reported altered acoustic transmission in dysplastic hips [7], [8], [9]. A recent study of DDH screening using sound transmission [10] suggested that transmitted energy was reduced in dysplastic hips, particularly in the 150–900 Hz range. In the current study, we investigated the utility of features extracted from the transfer function between left-to-right sound transmission for identifying normal and DDH patients.

Methodology: The study included 42 infants (32 females) with an age of 8±6 weeks (mean±SD). There dwere 28 normal subjects, 5 with unilateral DDH, and 9 bilateral DDH subjects (two with dysplasia of both

hips). During data collection, each patient was positioned supine with the knees and hips bent at about a 90-degree angle. Acoustic exciters were placed at the left and right anterior superior iliac spine (ASIS) and electronic stethoscopes were positioned at the greater trochanter of the left and right legs. Data were collected from the DDH clinic at NYU and is not publicly available. The inclusion criteria were subjects with suspected DDH (unilateral or bilateral) and normal controls without hip abnormalities as evaluated by physicians. Subjects with incomplete signal recordings or significant motion artifacts were excluded. The left-to-right anatomic asymmetry was assessed using the left-to-right transfer function (TFE) of transmitted signals. Figure 1 shows the TFE of a normal subject in Decibels, where higher TFE values indicate higher transmission asymmetry. Therefore, TFE features may be more useful for detecting unilateral DDH. Several TFE features were considered in the current study while attention was focussed on the 3 features that showed highest ability in separating the normal and DDH cases. The selected features were: the spectral crest, kurtosis and frequency with highest asymmetric transmission in the 200-800 Hz range. The crest ('mx') was calculated as the maximum TFE value after subtracting the average. This feature is known to estimate the relative "peakness" of TFE [11].

The three selected features are plotted in Figure 2, 3, 4 and a cutoff (dashed line) was manually chosen to separate the Normal from the DDH patients. Data points are labeled with the study subject numbers.

Results: Figure 2 shows the spectral crest and suggests that DDH patients were associated with higher crest values (which correspond to larger frequency-dependent acoustic transmission asymmetry). The subjects with dislocated hip joints are shown as solid red circles for both unilateral and bilateral cases. This figure also shows that with one feature (the crest) and a threshold value of 9.1dB, we can successfully identify DDH with a sensitivity of 92.9% (13/14). The specificity was low, i.e., 32.1% (9/28), with many normal cases having a crest above the threshold. Here, the unilateral and bilateral cases are combined as one DDH class when calculating sensitivity and specificity. The threshold was chosen such that unilateral dislocated hips were not misclassified and lie a reasonable distance above the threshold. Figure 2 also shows that one unilateral DDH subject (subject 45) was mis-identified as normal and another unilateral DDH subject

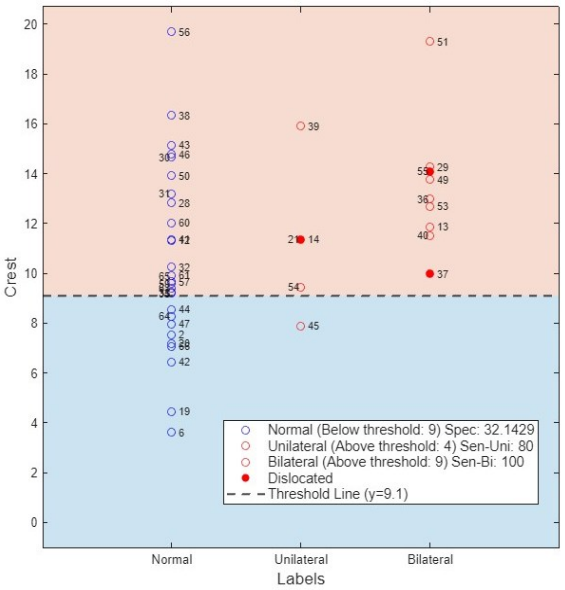

Figure 2. The spectral crest. A threshold is selected to separate the Normal from DDH cases (unilateral and bilateral). Data points are labeled with the study subject numbers. Open circles are DDH cases while solid circles are dislocation cases.

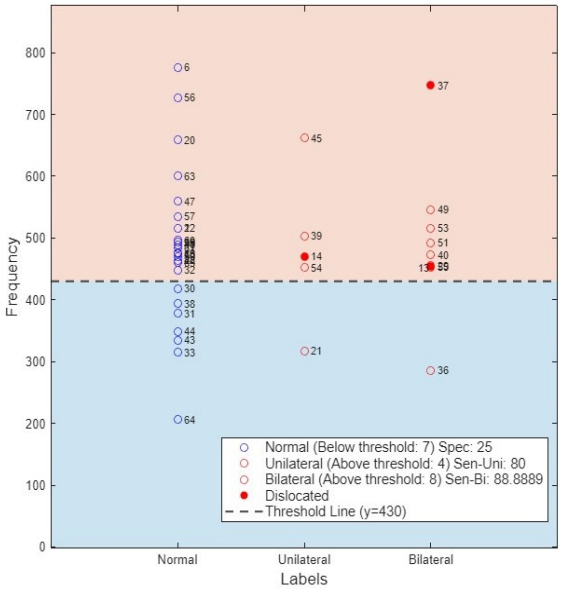

Figure 3. Class distribution for the frequency with maximum asymmetry. A threshold (dashed line) is used to separate normals and abnormals.

(subject 54) was close to the threshold line. To increase identification accuracy, additional features may further improve classification accuracy. These features were selected because they are extracted from TFE which capture statistical and frequency-domain aspects of acoustic transmission asymmetry.

Figure 3 shows the frequency with maximum asymmetry for all subjects. Here, we can see that subject 21 (unilateral DDH) and 36 (bilateral DDH) were misclassified. For a threshold of 430 Hz, the overall sensitivity was 85.7% (12/14) while the specificity was also low 25%, indicating a worse performance compared to the spectral crest of Figure 2. However, subject 45 has been detected accurately which was misclassified using the spectral crest. Figure 4 shows the kurtosis, where a threshold of 2.3 was chosen to detect unilateral cases that were missed. Here, Subject 45 was correctly identified with a noticeable margin from the threshold, but other unilateral (21, 54) and bilateral (37) subjects were misclassified. Here the specificity was 14.3% only.

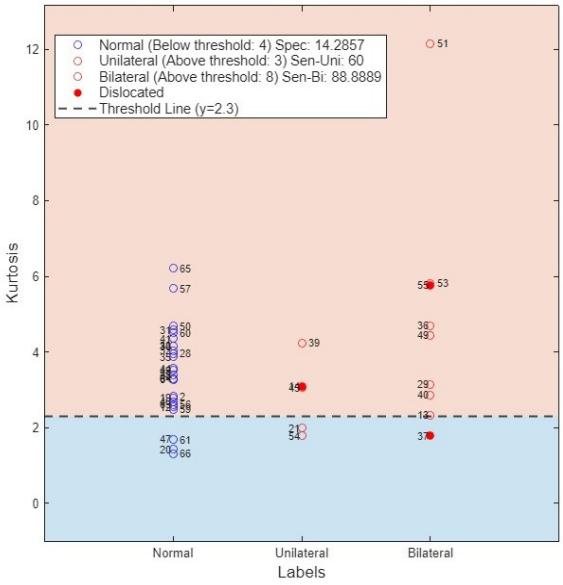

Figure 4. Class distribution for the spectral kurtosis with a threshold (dashed line) to separate normals and abnormals.

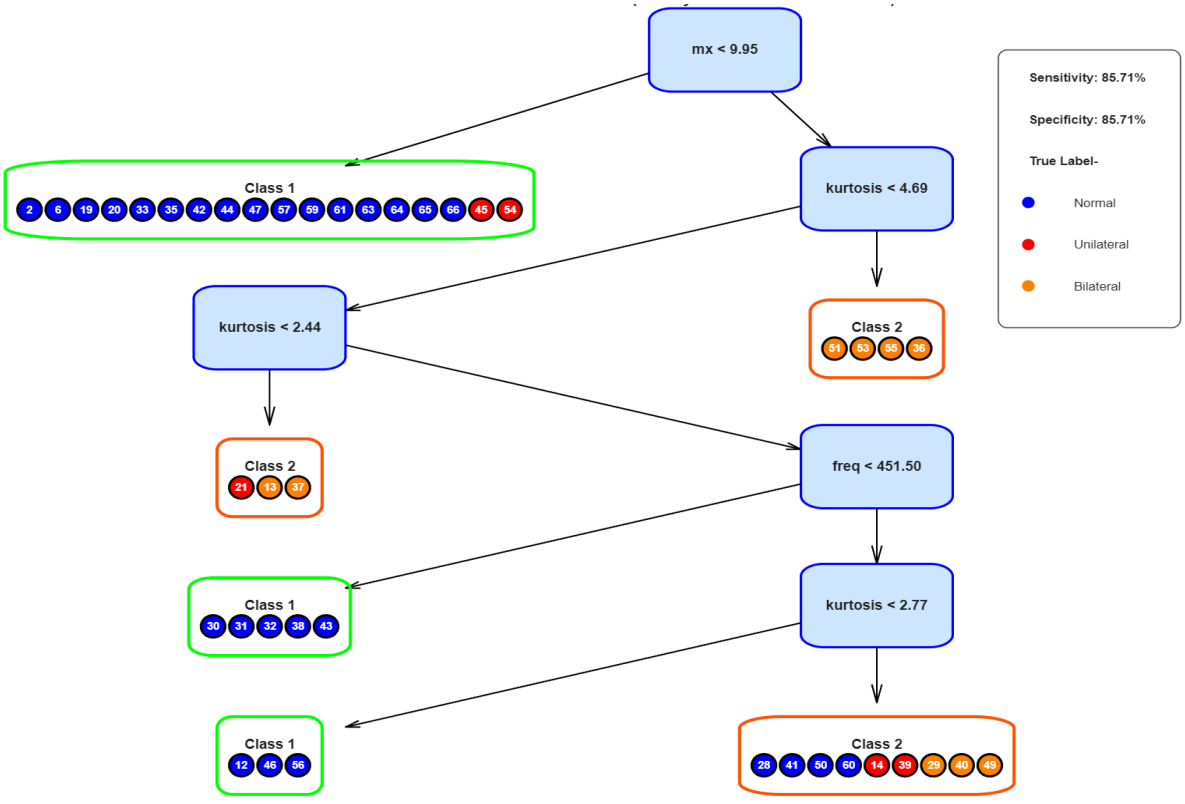

Figure 5. Decision tree classification with individual subject number and predicted class. Classification was based on a binary model (normal vs. DDH). Unilateral and bilateral cases (both are class 2) are labeled differently to provide more insight into the data distribution.

Table 1. The comparative results of the threshold method and decision tree

Method	Sensitivity (%)	Specificity (%)
Threshold with Crest (mx)	92.86	32.14
Threshold with Frequency	85.71	25.00
Threshold with Kurtosis	78.57	14.28
Decision Tree with default settings	85.71	85.71
Decision Tree with misclassifying cost 1.1	92.86	82.14

The results of Figure 2, 3 and 4 suggest that each feature by itself may not provide optimal performance and combining features may be beneficial. To investigate the utility of combining the features, a decision tree model was used with all the three features.

The decision tree model was trained and tested with all 42 subjects (Figure 5). Due to the small number of subjects, and for the purpose of initial feature exploration and to simplify the decision tree structure, the unilateral and bilateral classes were merged into one class (also called DDH class), resulting in a binary classification (normal vs abnormal/DDH). In Figure 6, we show the confusion matrix corresponding to the decision tree results in Figure 5, confirming the classification performance (sensitivity = 85.71%, specificity = 85.71%). This was implemented using the fitctree function of MATLAB with default settings. The decision tree is a powerful tool for elucidating the distribution of the dataset. It follows a tree-like structure that begins with a root node, which contains the entire dataset. It follows an iterative process that continues until the tree reaches the leaf nodes, which represent the final classification results. The structure provided clear insight into our features and their influence in making the classification decision. For more detailed interpretation of the classification results, the tree (Figure 5) displays the patient ID at each leaf node.

In Figure 5, different colors were used depending on the subject's actual class label. All subjects were used in training to explore how normal and abnormal subjects may be distributed in the feature space and to identify possible threshold values in the current study. Table 1 shows the results for all classification cases used (e.g., threshold on one feature and combing all features in a decision tree). The classification results of the decision tree showed a sensitivity was 85.7% and specificity was 85.7%. To maintain a sensitivity close to that of Figure 2, a cost of 1.1 was implemented in MATLAB and the resulting sensitivity and specificity were 92.86% and 82.14 %, respectively. Therefore, it can be concluded that combining the 3 features, using a decision tree approach, improved the classification performance in the current study. It is to be noted that these results provide potential performance metrics (since the same data was used for testing and training). When a larger data set is available in future studies, different data sets will be used for training and testing. In Figure 5, the spectral crest is the root node of the decision tree suggesting that it may be the most useful feature in distinguishing between normal and DDH subjects. Notably, the threshold of this node is 9.95 which was close to the cutoff value of 9.1 chosen in Figure 2. The decision tree results also indicate that kurtosis and frequency at the maximum asymmetry are useful TFE features that can increase specificity at comparable sensitivity values. Another important observation is that even with all three features utilized, it was not possible to separate subject 45 from the normal subjects.

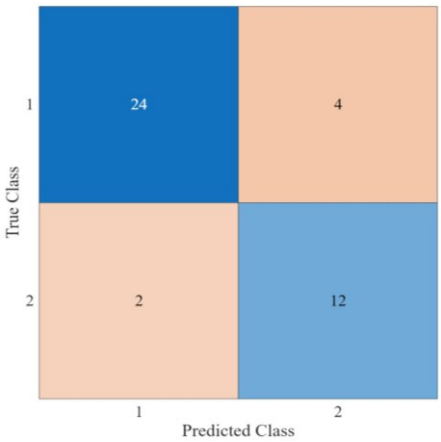

Figure 6. The confusion matrix for the decision tree with default settings.

Conclusion: The current study provides valuable insights into the distribution of our dataset and the characteristics of acoustic transmission for normal and DDH patients. A primary limitation of the study is the small dataset (42 subjects) with a small subset of DDH cases (5 unilateral, 9 bilateral). Another limitation is the nature of left-to-right transfer function which is mainly a measure of asymmetry. This

method was chosen both for ease of analysis, and because of the perceived higher incidence of unilateral compared to bilateral cases. The study suggested that bilateral DDH may also have high incidence, which increases interest in bilateral DDH detection. Interestingly, some bilateral cases exhibited measurable asymmetry, suggesting that complete symmetry between affected hips may be uncommon. Although these observations require validation with a larger dataset, the method discussed in this study has the potential in identifying both unilateral and bilateral abnormalities. Future studies need to involve more subjects and extract features that are beyond asymmetry such as input-output TFE. Future work will also investigate other classification methods to increase sensitivity, specificity and increase generality of the results.

ACKNOWLEDGMENTS

This publication is based on work supported by the National Science Foundation under Grant No. FW-HTF-P 2026516. The research reported in this publication was also supported by the National Heart and Lung Institute of the National Institutes of Health under award number R44 HL099053. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or National Institute of Health.

COI statement: Richard H. Sandler and Hansen A. Mansy are part owners of Biomedical Acoustics Research Company, which is the primary recipient of the NIH grant R44HL099053, as such they may benefit financially because of the outcomes of the research work reported in this publication.

REFERENCES

[1] S. K. Storer and D. L. Skaggs, “Developmental Dysplasia of the Hip,” in *Am Fam Physician,* 2006, vol. 74, no. 8, pp. 1310-1360.

[2] M. D. Sewell, K. Rosendahl, and D. M. Eastwood, “Developmental Dysplasia of the Hip,” in *BMJ*, 2009, vol. 339, no. nov24, pp. b4454–b4454, doi: 10.1136/bmj.b4454.

[3] A. Roposch, L. Q. Liu, F. Hefti, N. M. P. Clarke, and J. H. Wedge, “Standardized Diagnostic Criteria for Developmental Dysplasia of the Hip in Early Infancy,” in *Clinical Orthopaedics & Related Research,* Dec. 2011, vol. 469, no. 12, pp. 3451–3461, doi: 10.1007/s11999-011-2066-9.

[4] A. Vaquero-Picado, G. González-Morán, E. G. Garay, and L. Moraleda, “Developmental Dysplasia of the Hip: Update of Management,” in *EFORT Open Reviews*, Sep. 2019, vol. 4, no. 9, pp. 548–556, doi: 10.1302/2058-5241.4.180019

[5] C. Dezateux and K. Rosendahl, “Developmental Dysplasia of the Hip,” in *The Lancet*, 2007, vol. 369, Issue 9572, pp. 1541 - 1552.

[6] S. Scott Morey, “AAP Develops Guidelines for Early Detection of Dislocated Hips,” in *Am Fam Physician*, Feb 2001, pp. 565–566, 568, PMID: 11272302.

[7] K. S. C. Kwong, X. Huang, J. C. Y. Cheng, and J. H. Evans, “Acoustic Transmission in Normal Human Hips: Structural Testing of Joint Symmetry,” in *Medical Engineering & Physics*, Dec. 2003, vol. 25, no. 10, pp. 811–816, doi: 10.1016/s1350-4533(03)00113-9.

[8] T. Hassan et al., “An Acoustic Approach for Detection of Developmental Dysplasia of Hip,” in *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Dec. 2018, pp. 1–6. doi: 10.1109/spmb.2018.8615627.

[9] S. M. I. Kapicioglu and F. Korkusuz, “Diagnosis of Developmental Dislocation of the Hip by Sonospectrography,” in *Clinical Orthopaedics & Related Research*, Apr. 2008, vol. 466, no. 4pp. 802–808, doi: 10.1007/s11999-008-0163-1.

[10] T. Singh et al., "Detection of Hip Dysplasia in Infants Using Audible-Frequency Acoustic Transmission Measurements," in *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Dec. 2023, pp. 1–5. doi: 10.1109/SPMB59478.2023.10372804.

[11] G. Peeters, "A large set of audio features for sound description (similarity and classification) in the CUIDADO project," in *Tech. Rep*., IRCAM, Paris, France, *2004*. [Online]. Available: https://www.ircam.fr/.

[12] M. Chavoshi, G. Soltani, S. Shafiei Zargar, C. C. W. Wyles, H. M. Kremers, and P. Rouzrokh, "Diagnostic Performance of Clinical Examination Versus Ultrasonography in the Detection of Developmental Dysplasia of Hip: A Systematic Review and Meta-Analysis," *Archives of Bone and Joint Surgery*, May 2022, vol. 10, no. 5, pp. 403–412, doi: 10.22038/ABJS.2021.60504.2984.

[13] N. F. Woolacott, M. A. Puhan, J. Steurer, and J. Kleijnen, "Ultrasonography in screening for developmental dysplasia of the hip in newborns: Systematic review," in *BMJ*, Jun. 2005, vol. 330, no. 7505, p. 1413, doi: 10.1136/bmj.38450.646088.E0.

Using Acoustic Transmission Features and Decision Trees for Classification of Developmental Dysplasia of the Hip

M. Hossen, R. Sandler and H. Mansy

Biomedical Acoustics Research Lab, University of Central Florida, Orlando, Florida, USA

BIOMEDICAL ACOUSTICS RESEARCH LAB

College of Engineering and Computer Science — UCF

Abstract

- Developmental dysplasia of the hip (DDH) is a pediatric condition where the hip joint is improperly formed resulting in abnormalities in femoral head, acetabulum or both, which can lead to mild acetabular shallowness or complete dislocation [1]. DDH incidence may be as high as 4% to 6% in newborn infants[2], [3].
- Current newborn screening relies on physical examination (e.g., the Ortolani/Barlow tests). The Ortolani and Barlow maneuvers have low sensitivity (~36%) but high specificity (~98%) [5].
- One promising approach is the use of acoustic transmission to detect structural abnormalities in the hip joint. Several studies have reported altered acoustic transmission in dysplastic hips.
- In the current study, we investigated the utility of features extracted from the transfer function between left-to-right sound transmission for identifying normal and DDH patients.

Methodology

- The study included 42 infants with an age of 8±6 weeks.
- Normal: 28 subjects.
- DDH: 14 subjects (5 Unilateral, 9 Bilateral)
- During data collection, acoustic exciters were placed at the left and right anterior superior iliac spine (ASIS) and electronic stethoscopes were positioned at the greater trochanter of the left and right legs.
- The left-to-right anatomic asymmetry was assessed using the left-to-right transfer function (TFE) of transmitted signals.
- Figure 1 shows the TFE of a normal subject in Decibels.
- Features Extracted (200-800 Hz) include:
  1. Spectral Crest ('mx'), which is an indicator of relative "peakness" of the TFE.
  2. The frequency where the crest is located.
  3. Kurtosis, which is a Statistical measure of the TFE distribution broadness.

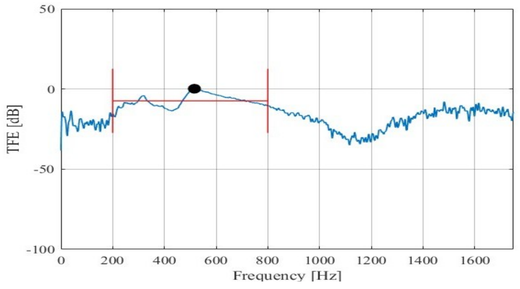

Figure 1. Left-to-right transfer function (TFE) of a normal subject. The horizontal red line represents the average value in 200-800 frequency range. Black dot is the maximum value within the range. The crest (aka. 'mx') is calculated as the maximum – average, in the selected frequency range.

Feature Performance (Threshold Method)

- Figure 2 shows the spectral crest and suggests that DDH patients were associated with higher crest values.
- With a threshold of crest=9.1dB, high sensitivity (92.9%) but low specificity (32.1%) can be achieved.
- The threshold was chosen such that the unilateral dislocated hip was not misclassified, while remaining at a reasonable distance above the cutoff.

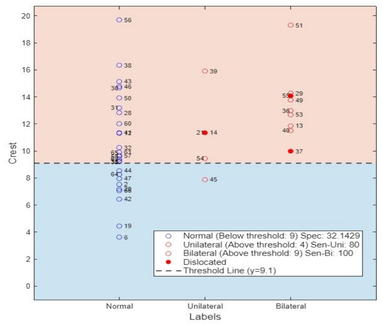

Figure 2. The spectral crest. A threshold is selected to separate the Normal from DDH cases (unilateral and bilateral). Data points are labeled with the study subject numbers. Open circles are DDH cases while solid circles are dislocation cases.

- Frequency is the second feature and yielded a sensitivity = 85.7% and specificity = 25% (Figure 3). Subjects #21 (unilateral DDH) and #36 (bilateral DDH) were misclassified.
- The sensitivity and, specificity for Kurtosis with 2.3 threshold were 78.6% and 14.3% respectively (Figure 4).

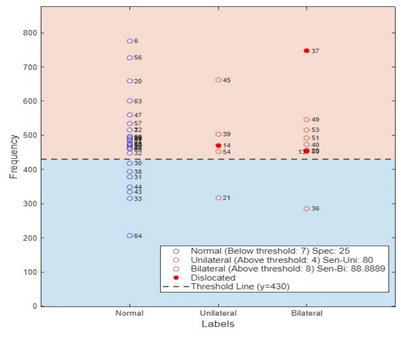

Figure 3. Class distribution for the frequency with maximum asymmetry. A threshold (dashed line) is used to separate normals and abnormals.

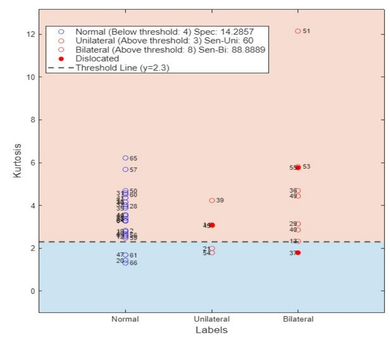

Figure 4. Class distribution for the spectral kurtosis with a threshold (dashed line) to separate normals and abnormals.

Decision Tree Classification

- To investigate the utility of combining the three features, a decision tree model was constructed.
- The decision tree model was trained and tested with all 42 subjects (Figure 5).
- To simplify the decision tree structure, the unilateral and bilateral classes were merged into one class, resulting in a binary classification.
- The structure provided insight into feature utility and their influence in making the classification decision.

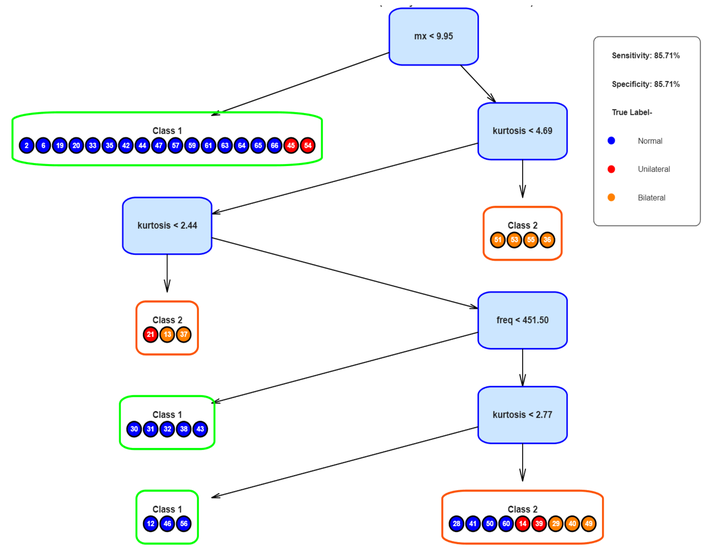

Figure 5. Decision tree classification with individual subject number and predicted class. Classification was based on a binary model (normal vs. DDH). Unilateral and bilateral cases (both are class 2) are labeled differently to provide more insight into the data distribution.

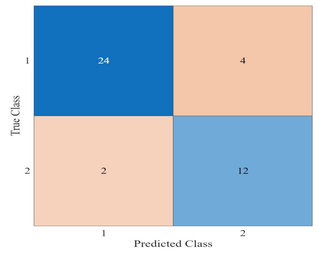

Figure 6. The confusion matrix for the decision tree with default settings.

Discussion

- Table 1 shows the results for all classification methods studied (single threshold and decision tree).
- The classification results of the decision tree showed a sensitivity of 85.7% and specificity of 85.7%.
- The decision tree model provided higher sensitivity while maintaining the higher specificity compared to the single feature threshold methods.
- To maintain a sensitivity close to that of Figure 2, a cost of 1.1 was implemented which raised the sensitivity of the decision tree model to 92.9%. The corresponding specificity was 82.1%

Table 1. The comparative results of the threshold method and decision tree

Method	Sensitivity (%)	Specificity (%)
Threshold with Crest (mx)	92.86	32.14
Threshold with Frequency	85.71	25.00
Threshold with Kurtosis	78.57	14.28
Decision Tree with default settings	85.71	85.71
Decision Tree with misclassifying cost 1.1	92.86	82.14

Conclusion

- The current study provides valuable insights into the distribution of our dataset and the characteristics of acoustic transmission for normal and DDH patients.
- A primary limitation of the study is the small dataset (42 subjects) with a small subset of DDH cases (5 unilateral, 9 bilateral). Another limitation is the nature of left-to-right transfer function which is mainly a measure of asymmetry.
- Interestingly, some bilateral cases exhibited measurable asymmetry, suggesting that complete symmetry between affected hips may be uncommon.
- Although these observations require validation with a larger dataset, the method discussed in this study has the potential in identifying both unilateral and bilateral abnormalities.
- Future studies need to involve more subjects and extract features that are beyond asymmetry.
- Future work will also investigate other classification methods to increase sensitivity, specificity and increase generality of the results.

References

[1] S. K. Storer and D. L. Skaggs, "Developmental Dysplasia of the Hip," in Am Fam Physician, 2006, vol. 74, no. 8, pp. 1310-1360.

[2] M. D. Sewell, K. Rosendahl, and D. M. Eastwood, "Developmental Dysplasia of the Hip," in BMJ, 2009, vol. 339, no. nov24, pp. b4454–b4454, doi: 10.1136/bmj.b4454.

[3] A. Roposch, L. Q. Liu, F. Hefti, N. M. P. Clarke, and J. H. Wedge, "Standardized Diagnostic Criteria for Developmental Dysplasia of the Hip in Early Infancy," in Clinical Orthopaedics & Related Research, Dec. 2011, vol. 469, no. 12, pp. 3451–3461, doi: 10.1007/s11999-011-2066-9.

[4] S. Scott Morey, "AAP Develops Guidelines for Early Detection of Dislocated Hips," in Am Fam Physician, Feb 2001, pp. 565–566, 568, PMID: 11272302.

[5] M. Chavoshi, G. Soltani, S. Shafiei Zargar, C. C. W. Wyles, H. M. Kremers, and P. Rouzrokh,"Diagnostic Performance of Clinical Examination Versus Ultrasonography in the Detection of Developmental Dysplasia of Hip: A Systematic Review and Meta-Analysis," Archives of Bone andJoint Surgery, May 2022, vol. 10, no. 5, pp. 403–412.