ANALYSIS OF FINGERPRINT RECOGNITION PERFORMANCE ON INFANTS

by

Samuel Reiff

A Thesis

Submitted to the Faculty of Purdue University
In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Technology Leadership & Innovation
West Lafayette, Indiana
August 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Stephen Elliott, Chair

Polytechnic Institute

Dr. J. Eric Dietz

Polytechnic Institute

Kevin O'Connor

International Center for Biometric Research (ICBR), Purdue University

Approved by:

Dr. Nathan Mentzer

I dedicate this to my friends and family, and to all that have positively impacted my life. To those that helped shape the man I have become. To my parents for keeping me accountable, my siblings for being motivation to succeed, and all my friends that have supported me and encouraged me throughout this journey. May I impact others' lives as you all have impacted mine.

ACKNOWLEDGMENTS

This has been a journey, and they say it takes a whole village. In this case, there have been many that have supported me throughout this challenging and rewarding journey. First, I would like to thank my Lord and Savior, Jesus Christ. He has provided me with many blessings and has allowed me to make the most of what I have been given. He has never let me down and opened doors I never thought were possible. I hope that I can continue to work and live in a way that makes You proud.

I would like to thank my advisor and committee chair, Dr Stephen Elliott. I am grateful for the academic and professional opportunities and support you have provided. You were an integral part of my academic success, and though I maybe did not like to admit it at the time, your feedback strengthened my research and academic/professional preparedness. You have been an insightful mentor that has pushed me to work hard, and I have appreciated the knowledge you have bestowed upon me.

Kevin O'Connor, you have also been an integral part of my success. You have provided sound advice and critical feedback, and I really enjoyed teaching with you. You exhibited a good blend of willingness to listen/work with students, and accountable feedback. Dr. Dietz, you have also been a positive influence on my academic success. Your passion for teaching and compassion for students was contagious. You helped bring my academic/professional interests into focus. Thank you to Professors Jacob, Denton, and Cheung. Even though your influence occurred during my undergraduate years, you helped shape the type of student I am and the type of employee I want to be. Though you did not expect perfection, you demanded excellence and hard work. You positively impacted my desire and success while pursuing my master's degree,

and those lessons you taught me will continue to drive me as I continue to move forward. You taught me how to learn and how to enjoy school. I will forever have a learning mind and heart.

Thank you, to all parents or guardians of the infants that participated in this study. Without you, there would be no study and without this study, this would not be my research. The field of infant biometrics is still very small, so your contribution made a huge impact.

I would like to thank my parents, Park and Melissa. I love you and appreciate all you have done for me. Mom, you have had the greatest influence on my life, academically. You taught me that the best teachers are the ones who learn along with you. You have instilled in me a love for learning, a love for teaching, a love for others, and an open-mindedness to listen to others. Thanks to this, I have been able to blossom as an individual and as a member of a team, whether it is in a classroom or in society. I also owe my passion and fire to you, that drive has been integral to me pursuing a graduate degree and persevering through all my academic career. It is amazing to see how far I have come, and I owe a lot of that to you and your support. Dad, you have pushed me to be a man of integrity, wisdom, and one who values accountability and hard work. These values have helped me achieve what I would not have been able to otherwise. You also helped me develop the analytical mindset that got me through the technical curriculum of my college career, so thank you.

I would like to thank my brothers and sisters, all six of them! You have given me much to be grateful for. I strive to succeed for myself, but I also want to be a positive example for you all. Your love has gotten me through hard times and your presence has driven me to aim higher and work harder. To my grandparents, thank you for teaching me that nothing in life is given, you must earn it. Thank you for showing me how to work hard to achieve my goals.

To Brian Biedenbach, thank you for being a beacon of spiritual support. By living by example, you greatly influenced my life. Apart from my family, you have had the biggest hand in shaping the man I am today. To Jennie Thomas, thank you for supporting me at my lowest of lows. Thank you for seeing the potential I have always had, and for expecting nothing but the best from me. You always knew I was more than capable. To Doug and Sally Stitzer, thank you for taking me under your wing and being intentional about being there for me. You gave me hope at times where I could not seem to find it on my own. Your kindness and wisdom are much appreciated. To my other mentors, I have not forgotten you. Your support and wisdom live on and continue to help me every day.

To my friends who have been with me for each step of this academic journey, thank you. You know who you are. I appreciate the old friends that have loved, encouraged, supported me. You all gave me hope when I was at my weakest or when I was starting to doubt myself. Words cannot explain my appreciate for you. I also appreciate the new friends, who have entered the picture during this journey. You have all come into my life at the precise moments I needed support, and though I never asked for it or expected it, you gave it freely and unconditionally. So, thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEFINITIONS

Biometric: “A measurable, physical characteristic or biological characteristic used to recognize the identity or verify these claimed identity of an enrollee” (Association of Biometrics, 1999, p.2).

Biometric Aging: “The gradual decrease in a system performance caused by the changes suffered by the users’ trait in the long-term (which cannot be avoided as is inherent to human nature)” (Lanitis, 2010, p.32).

Biometric non-mated comparison trials: They “have historically been referred to as ‘impostor trials’” (ISO / IEC JTC 1 SC 37, 2005, p.19).

Chameleon: “A person who is a chameleon matches well in general, both to themselves and to others. They are likely to cause false accepts but not false rejects” (Teli et al., 2011, p.6).

Detection error trade-off curve (DET curve): A “modified ROC curve that plots error rates on both axes (false positives on the x-axis and false negatives on the y-axis)” (ISO / IEC JTC 1 SC 37, 2005, p.7).

Dove: “A person who is a dove matches very well against themselves and poorly against others” (Teli et al., 2011, p.6).

Failure to acquire rate (FTAR): The “proportion of a specified set of biometric acquisition processes that were failures to acquire” (ISO / IEC JTC 1 SC 37, 2005, p.20).

Failure to enrol rate (FTER): The “proportion of a specified set of biometric enrolment transactions that resulted in a failure to enroll” (ISO / IEC JTC 1 SC 37, 2005, p.5).

False match rate (FMR): The “proportion of zero-effort impostor attempt sample features falsely declared to match the compared non-self” (ISO / IEC JTC 1 SC 37, 2005, p.5).

False non-match rate (FNMR): The “proportion of genuine attempt sample features falsely declared not to match the template of the same characteristic from the same user supplying the sample” (ISO / IEC JTC 1 SC 37, 2005, p.5).

Genuine attempt: A “single good-faith attempt by a user to match their own stored template” (ISO / IEC JTC 1 SC 37, 2005, p.2).

Habituation: The “degree of familiarity of a biometric capture subject with the biometric capture process” (ISO / IEC JTC 1 SC 37, 2017, p.14).

Identification: The “process of searching against a biometric enrolment database to find and return the biometric reference identifier(s) attributable to a single individual (ISO / IEC JTC 1 SC 37, 2005, p.18).

Impostor attempt: An “attempt of an individual to match the stored template of a different individual by presenting a simulated or reproduced biometric sample or by intentionally modifying his/her own biometric characteristics” (ISO / IEC JTC 1 SC 37, 2005, p.3).

Matching score: “Measure of the similarity between features derived from a sample and a stored template or a measure of how well these features fit a user’s reference model” (ISO / IEC JTC 1 SC 37, 2005, p.2).

Phantom: “A person who is a phantom matches poorly in general, both to themselves and to others. They are likely to cause false rejects but not false accepts” (Teli et al., 2011, p.6).

Receiver operating characteristic curve (ROC curve): A “plot of the rate of “false positives” (i.e., impostor attempts accepted) on the x-axis against the corresponding rate of “true positives”” (ISO / IEC JTC 1 SC 37, 2005, p.6).

Template: A “user’s stored reference measure based on features extracted from enrollment samples” (ISO / IEC JTC 1 SC 37, 2005, p.2).

Template Aging: “Occurs when the quality of the match between an enrolled biometric sample and a sample to be verified degrades with increased elapsed time between the two samples” (Fenker & Bowyer, 2011).

Verification: The “application in which the user makes a positive claim to an identity, features derived from the submitted sample biometric measure are compared to the enrolled template for the claimed identity, and an accept or reject decision regarding the identity claim is returned” (ISO / IEC JTC 1 SC 37, 2005, p.4).

Worm: “A person who is a worm matches themselves poorly and other people relatively well. They result in a disproportionate number of errors, both false rejects and false accepts” (Teli et al., 2011, p.6).

LIST OF ABBREVIATIONS

DET Curve: Detection error tradeoff curve

EER: Equal error rate

FMR: False match rate

FNMR: False non-match rate

FTA: Failure to acquire

FTX: Failure to extract

ROC Curve: Receiver operating characteristic curve

SSI: Stability Score Index

ABSTRACT

In this study, any change in fingerprint performance, image quality and minutiae count for infants in three different age groups was evaluated (0-6, 7-12, and >12 months). This was done to determine whether there is a difference in performance between infant age groups for a fingerprint recognition system.

The purpose of this research was to determine whether there is a difference in infant fingerprint performance and image quality metrics, between three different age groups (0-6, 7-12, and >12 months old), using the same optical sensor? The data used for this secondary analysis was collected as part of a longitudinal multimodal infant study, using the Digital Persona U.are.U 4500. DET curves, zoo analysis, and image quality metrics were used to evaluate performance and quality factored by infant age group.

This study found that there was a difference in image quality and minutiae count, genuine and impostor match scores, and performance error rates (EER) between the three age groups. Therefore, quality and performance were dependent on age. While there was a difference in performance between age groups, there was generally stability for subjects who overlapped between multiple age groups. Difference in performance was most likely due to the difference in physical characteristics between subjects in each age group, rather than individual instability. The results showed that it could potentially be feasible to use fingerprint recognition for children over the age of 12 months.

CHAPTER 1. INTRODUCTION

1.1 Introduction

According to Cummins & Midlo (1961), biological evidence shows that fingerprints are fully formed after six months of fetal life, thus are fully developed at birth. This points toward fingerprint as a usable modality for infant biometrics. However, the quality of an infant's fingerprint tends to be very poor (Jain et al., 2016). The question then revolves around how infants perform and in what ways can that process be improved. When dealing with infants, their natural behavior does not lend itself well to collecting biometric data. Though, if the infant's physical fingerprints are fully developed, then it may be possible with more testing to develop a sensor or technology for this population.

According to Muramoto (2015), it may be possible to distinguish between a latent fingerprint that is one day or one week old, one week or one month old, one month and a couple of months old, and so on. It could then also be possible to distinguish live scan fingerprints and their biometric performance between these age groups. Certain age groups may be better or worse performers, and if so, this knowledge could potentially be used a tool to guide researchers toward a subset of infants where fingerprint recognition or a technology can be adequately used.

This chapter serves to introduce the topic of interest and lay the groundwork for the rest of this study. This first chapter can be used as a guide to better understand the rest of this thesis by including the significance of the problem, statement of purpose, scope, research question, assumptions, limitations, and delimitations.

1.2 Statement of the Problem

Infant biometrics are still largely in a research phase with a lot still left to be discovered, as large datasets on infants are very difficult to obtain (Jain et al., 2016). Fingerprint recognition as it stands today is commonly viewed as unviable for use with the infant population (Jain et al., 2016). This creates a knowledge gap regarding the cause of poor infant fingerprint performance. Poor performance could unavoidable, but maybe it is the device being used, or maybe there are specific infant age groups that consistently perform better or worse. This prods researchers to ask the following question: “for infants, does the particular sensor matter?”, and “does their specific age matter?”

1.3 Research Question/Hypotheses

The research question being addressed is as follows: is there a difference in infant fingerprint performance and image quality metrics, between different age groups (0-6 months, 7-12 months, and >12 months old), using the same optical sensor?

Metrics used to evaluate include: Image quality and minutiae count, genuine and impostor match scores, detection error tradeoff (DET) curves and equal error rates (EER), and zoo analysis.

Sub Questions:

- Is there a difference in performance between infant age groups?
    - Metrics include: FTX rates, Genuine & Impostor Match Scores, DET Curves & EER, Zoo Analysis
- Is there a difference in image quality metrics between infant age groups?
    - Metrics include: Image Quality & Minutiae Count

- Is there individual stability in performance for infants who aged into a different age group during the longitudinal data collection?
    - Metrics include: Zoo Analysis, Stability Score Index

The hypotheses were as follows:

There is no difference in infant fingerprint performance between the three age groups

- H_o: $\mu_1 = \mu_2 = \mu_3$

  H_a: $\mu_1 \neq \mu_2 \neq \mu_3$

  Where μ_1 is the 0-6 months old group, μ_2 is the 7-12 months old group, and μ_3 is the >12 months old group.

There is no difference in infant fingerprint image quality metrics between the three age groups

- H_o: $\mu_1 = \mu_2 = \mu_3$

  H_a: $\mu_1 \neq \mu_2 \neq \mu_3$

  Where μ_1 is the 0-6 months old group, μ_2 is the 7-12 months old group, and μ_3 is the >12 months old group.

There is stability in individual infant fingerprint performance for those who aged into a different group during the longitudinal data collection

- H_o: $\mu_1 = \mu_2 = \mu_3$

  H_a: $\mu_1 \neq \mu_2 \neq \mu_3$

  Where μ_1 is the 0-6 months old group, μ_2 is the 7-12 months old group, and μ_3 is the >12 months old group.

1.4 Significance of the Problem

According to Bharadwaj et al. (2010), Corby et al. (2006), Jain et al. (2014), Jain et al. (2016), Jain et al. (2017), Jia et al. (2010), and Jia et al. (2012), clean and precise infant biometrics are extremely difficult to obtain due to an infants' rapidly changing physical growth and development, their natural unintentionally uncooperative behavior, and the fact that most biometric systems are designed with a population of adults in mind. Additionally, there are a lot of factors or variables to consider within a biometric system. These factors, caused by individual (their fingerprint) and the machine (algorithm/template aging), may each affect quality and performance. Age is just one of them, and there currently is no one method that is sufficient for use in infant fingerprint recognition (Jain et al., 2014).

According to Jain et al. (2014), Jain et al. (2016), and Jain et al. (2017), the quality of captured fingerprint images tends to be very poor in infants. This is mostly due to the difficulty of getting infants to properly interact with a fingerprint sensor. Infants' fingers get wet from them putting their fingers in their mouth, they often time clench and unclench their fists and move their hands and fingers in seemingly random patterns, can be balled up (Jain et al., 2014). Additionally, the spacing of ridges and valleys in infants' fingerprints is smaller and more compact than that of an adult fingerprint (Jain et al., 2014). This creates an additional step in feature extraction, accounting for this spacing discrepancy (Jain et al., 2014). This research looked at whether the difference in fingerprints due to specific age ranges accounted for any of these discrepancies and impacted quality or performance.

There has recently been an emphasis by researchers, such as those from Bharadwaj et al. (2010), Corby et al. (2006), Jain et al. (2014), Jain et al. (2016), Jain et al. (2017), Jia et al. (2010), and Jia et al. (2012), to study the effect that age has on biometrics (specifically with infants) and how to mitigate some issues that age presents. In addition to the effect that age and infants have on biometric performance, there has also been a larger emphasis on being able to non-intrusively identify or verify a child’s identity to combat child and infant trafficking (UNICEF USA, 2019). According to UNICEF USA (2019), child trafficking occurs in all 50 U.S. states, and 1 in 4 victims of all trafficking are children. In many places, including Nigeria, infant trafficking has been growing since 2006 (Makinde, 2015). With a successful and proven method of identifying/verifying infants using fingerprint recognition, child trafficking could be better protected against and it could add to the knowledge and ability needed to develop more accurate biometric systems in general.

1.5 Statement of Purpose/Scope

The research questions this study attempted to answer are as follows: “is there a difference in infant fingerprint performance and image quality metrics, between different age groups (0-6 months, 7-12 months, and >12 months old), using the same optical sensor? Additionally, is there stability in individual infant fingerprint performance across age groups for children who aged into a separate age group during the longitudinal data collection?”

Quantifying the change that can occur in a fingerprint recognition system is important, as biometrics are becoming used more frequently for authentication purposes (identification and verification) as time passes. An “ideal” biometric should contain permanence (Jain et al., 2002). However, people are known to age and physically change over time, and aging is a factor that can be difficult to properly evaluate due to a compromise of factors when developing biometric systems, or the lack of a sufficient sample size. Time is just one variable that affects fingerprint recognition systems. There are many other variables such as the type of sensor technology, type of interaction, the force being applied by the individual, and the environment data is being captured in that affects fingerprint performance. There are other obvious circumstances that could affect a user's ability to perform well on a fingerprint scanner. These include an individual injuring their hand or losing a finger, or physical deformities at birth. There are also more subtle occurrences, such as a decrease in fingerprint definition due to wear and tear or a decrease in elasticity (Lanitis, 2010). However, time and more specifically, aging, is an area of research where much is still left to be uncovered.

This research conducted secondary analysis to establish how much change in performance and image quality for infant fingerprints between specific infant age groups (0-6, 7-12, and >12 months old). Data were collected using the Digital Persona U.are.U 4500 fingerprint sensor. This dataset was chosen for this secondary analysis because it is taken from the same multimodal infant biometric data collection that Hutchison (2018) used for their secondary analysis on infant iris recognition. This study builds upon the findings of that study and adds to the general body of knowledge in infant biometrics by focusing on fingerprint recognition. Since the variable of interest is age, any differences in performance or image quality would be attributed to the variable of infant age.

1.6 Assumptions

Assumptions in this study include:

1. The data was properly collected and was not compromised or inaccurate.

1.7 Limitations

This study was limited by the following.

1. The infants may have been uncooperative, which could have affected the quality of the collected data.

2. This study analyzes fingerprint recognition in only three specific age groups (0-6, 7-12, and >12 months old).
3. This was a secondary data analysis.

1.8 Delimitations

This project's delimitations include the following.

1. Fingerprint was the only modality investigated in this study, all other modalities (iris, face, etc.) were not evaluated.
2. The correlation between fingerprint image quality metrics and an individual's matching performance will not be evaluated.
3. Only one fingerprint sensor and matching algorithm was used in this study The performance of different sensors or algorithms were not investigated, as the affect these would have on performance was beyond the scope of this study. Infant behavior was beyond the scope of this study.

CHAPTER 2. REVIEW OF LITERATURE

This study evaluated the performance and image quality of infant fingerprints. The literature review was divided into seven main subsections: biometrics and authentication, the basics of fingerprint recognition, performance and evaluation metrics, aging, general infant biometrics, infant fingerprint recognition, and challenges with collecting infant data.

2.1 Biometrics and Authentication

In biometrics, people are identified by behavioral and physiological attributes. Physiological characteristics consist of unique traits that individuals were born with. Behavioral characteristics are traits that are developed over time, such as writing a signature. Biometrics are used for authentication purposes, evaluating these physiological and behavioral traits that belong to a given individual. Biometrics are separated into modalities. A biometric modality is a category by which a biometric system is classified based on the human trait that acts as the data input (Jagadeesan, 2010). Some common modalities that are or have been proposed for implementeion include voice, fingerprint, iris, face, voice, ear, gait, keystroke dynamics, signature, palm, and hand geometry. Each modality possesses strengths and weaknesses and could be more beneficial to use than other modalities depending on the specific application. Furthermore, ease of use, throughput, and other additional factors also contribute to the selection of a biometric modality. For example, if a facility storing classified information needed to be secured, the accurate protection of the sensitive information would be prioritized before achieving faster throughput. If the use-case only required access control for a public location, such as office or apartment buildings (Senseon Secure Access, 2020), high-efficiency throughput might be prioritized over maximum security.

Authentication has been described as the binding of an identity to a subject (Bishop, 2003) and can be performed using various methods. Authentication can be performed based upon what one has, what one is, or what one knows. What one knows would consist of intangible codes, such as passwords and PINs (personal identification numbers). What someone has, or a token, would include passports, badges, and ID cards. With regards to access control, someone might have a key or a card that grants them access to a room. These two methods of identification "are not able to meet the growing demands for stringent security in applications such as national ID cards, border crossings, government benefits, and access control" (Jain & Kumar, 2010). As a result, biometric recognition is being developed to adjust and adapt to the rapidly growing quantity of applications used for authentication. Biometrics are used to analyze and determine who someone is based upon traits that are specific to them, such as fingerprint, face, and iris recognition. Sometimes more than one biometric modality can be used at the same time. This would be two-factor authentication. An example of two-factor authentication is a hand geometry device. Hand geometry combines a biometric and a PIN (or magnetic stripe, depending on the implementation method), that is linked to the specific individual. Biometrics have been shown to be 'universal', as a good quality biometric sample can be acquired from all but two percent of the population, due to constraints such as disabilities and scars (Jain, 2005). According to Zibran (2012), while biometric traits are 'universal', they are not always 'invariant'. This adds another ripple to the accuracy and security conundrum of a biometric system.

The general biometric model, displayed below in *Figure 2.1*, provides a comprehensive overview of the authentication process for a generic biometric trait and is a valid for any biometric modality. There are five important stages that are identified: data capture, signal processing, data storage, matching, and decision making. Enrolment is the initial presentation of the biometric trait, the template creation, and its storage. There are two modes of performance, verification, and identification. Verification is a 1:1 comparison, verifying that a user is who they claim to be. Identification is a 1:N comparison, identifying a specific individual amongst a population of individuals (ISO / IEC 2382-37, 2017).

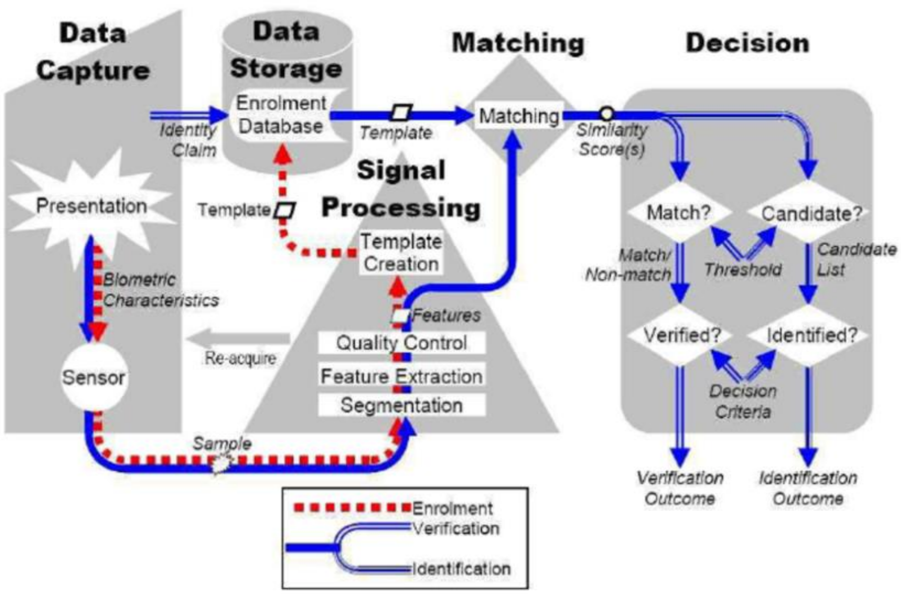

Figure 2.1 General Biometric Model (ISO/IEC TR 24741, 2006)

When the authentication process is initiated, for either verification or identification, the user presents their biometric trait to the sensor again and this new presentation is compared to the stored template image. This is the matching portion of the general biometric model. A similarity score is produced, evaluating how similar those two images are to each other. A genuine match score is the result of a user compared against themselves. An impostor match score is the result of a user compared against anyone else. *Figure 2.2* illustrates an example of an estimated probability density of this distribution. Depending upon the system's threshold, it is determined whether the user is who they say they are. This is the decision-making portion of the general model.

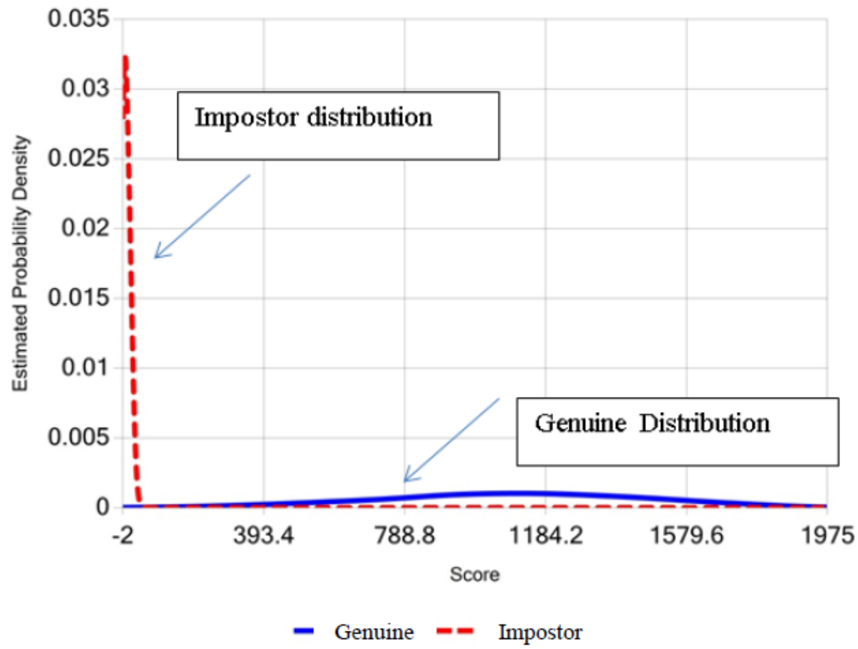

Figure 2.2 Distribution of Impostor and Genuine Scores

The distribution of genuine and impostor match scores shown above is used to evaluate a biometric system's performance level. The similarity score is produced by the matching algorithm and determines how similar the new sample presented for identification/verification is to the stored template. The threshold can be set, appropriate to a specific biometric system, based on these distributions. An example of this would be an individual using their fingerprint to unlock their phone. If the biometric threshold is set too high, then the genuine user could be denied access. This would likely cause frustration since the device belongs to them. However, if the threshold is set to low, then an impostor attempt could be granted access. This could potentially be much worse since annoyance is being traded for a lack of security. Analysis of biometric performance allows for the proper threshold to be set, according to the application and level of convenience/security that is required. Ground-truthing is used to verify that an individual is truly being matched against themselves, and correctly produces high genuine and low impostor match scores. This provides a system check for the system to ensure accurate and precise results. However, ground-truthing is performed by the researcher or data collector rather than a computer. There are challenges to this approach since the process is only as accurate as the individual is.

2.2 Basics of Fingerprint Recognition

The most common biometric modality employed worldwide for authentication purposes is the fingerprint (Violino, 2015). Fingerprint sensors are in phones and laptops and are the most common biometric in law enforcement and border control. Good and consistent performance matters when using technology of all kinds, and biometrics are no different. Border safety and municipal security, such as face recognition cameras in police stations or airports, are important for the well-being of those that live in that country or city. Additionally, the privacy and security of personal information that is stored on a personal phone or computer are also important.

Regarding the previously explained general biometric model, using fingerprint recognition as an example, the first part is the enrollment process. It is generally assumed that the same fingerprint sensor used for the enrollment process is used for the verification process (AlShehri, Hussain, AboAlSamh & AlZuair, 2018). If a different sensor is used, some biometric performance can be more exposed to technological and environmental factors. An individual presents their finger to the sensor, and the image is captured. After the image is successfully captured the features of the individual’s fingerprint are extracted and analyzed, and it is determined whether image quality expectations have been met. There are three common types of fingerprint matching algorithm: minutiae based, pattern-based, and hybrid. A matching algorithm can either be live-scan or latent. Live-scan refers to a live presentation of the fingerprint and latent refers to a fingerprint that has been left behind. The algorithm used for this study was live-scan and minutiae-based. For a minutiae-based matching algorithm extracts minutiae points from a fingerprint. These fingerprint features include ridge endings, bifurcations, deltas, and cores. If quality expectations have not been met then no information was extracted from the fingerprint, resulting in a failure to acquire (FTA). If the image passes the quality check, it is stored as that user's template. The template is based on biometric features extracted from enrollment samples (ISO / IEC JTC 1 SC 37, 2017). "Biometric templates store a coded representation of specific personal characteristics of a subject that can be used for personal identification" (Lanitis, 2010, p. 1). A template is the enrolment image that is captured upon an individual's first presentation to the biometric system. This template is then stored in the system's database and is compared to the captured image from all future presentations claiming to belong to that user.

Of the 150 different local ridge characteristics, two of the most prominent ridge characteristics that minutiae-based algorithms look for are ridge endings and ridge bifurcations. (Hong, 1998, p. 777). An ending is where a ridge ends abruptly, and a bifurcation is where one ridge branches or splits into two ridges (Hong, 1998, p. 777). Flat captures refer to a user presenting their fingerprint by pressing it flat against the sensor's scanning area. Rolled captures refer to a user placing one side of their finger against the sensor's scanning area and rolling it across to the other side of their finger. Flat fingerprints can generally produce up to 40 minutiae points, whereas rolled fingerprints can usually generate up to 140 minutiae points (Hong, 1998, p. 777). Larger scanning areas allow the sensor to capture more minutiae, and rolling a fingerprint creates more minutiae points for the sensor to pick up (Hong, 1998). For minutiae-based algorithms, this could potentially present a problem as missing minutiae or extra minutiae could negatively impact the matching algorithm’s ability to correctly identify an individual. These algorithms not only have to identify the correct points within a fingerprint and correctly match them to the right fingerprints, but they also need to be able to differentiate the fingerprint minutiae from smudges around the edges of the scanning area that are not part of the finger being presented. This can result in added minutiae points being marked, thus affecting the algorithm’s ability to properly match fingerprints (Hong, 1998).

The presentation of the fingerprint could also affect the quality of the captured image due to uneven force or pressure being applied to the sensor’s scanning area. When pressed too hard, the ridges can fold over and appear dark, as they overlap the area occupied by the valleys (Mason, 2014, p. 587). Too little pressure being applied can lead to the ridges being too light on the fingerprint image. This makes identifying specific ridge characteristics, or minutiae points, extremely hard (Mason, 2014, p. 587). This not only contributes to poor image quality, but it also creates problems for minutiae based-matching algorithms. According to Petrelli (2009), force significantly impacts performance on a ten-print fingerprint sensor. “For image quality among the thumbs, the highest quality images were gathered from the 6N – 12N force level”, and “the recommended force range to collect thumb images from a subject is 6N-12N” (Petrelli, 2009, p. 101). For the other four fingers, “higher quality images occur when a subject exerts from 12N-20N” (Petrelli, 2009, p. 101). Additionally, rolling fingerprints can create smudges due to uneven force being applied throughout the rotation of the finger. The finger can slide and create a blur in the image which negatively impacts the image quality of the captured fingerprint (Alonso-Fernandez, 2006, p. 423). In *Figure 2.3*, Khan (2015) shows a good quality fingerprint, a medium quality fingerprint, and a poor-quality fingerprint, from left to right. *Figure 2.4* shows different types of fingerprint pattern classes as well as commonly identified minutiae points of a fingerprint.

Figure 2.3 Examples of Good, Medium, and Poor Quality Fingerprint Captures (Khan, 2015)

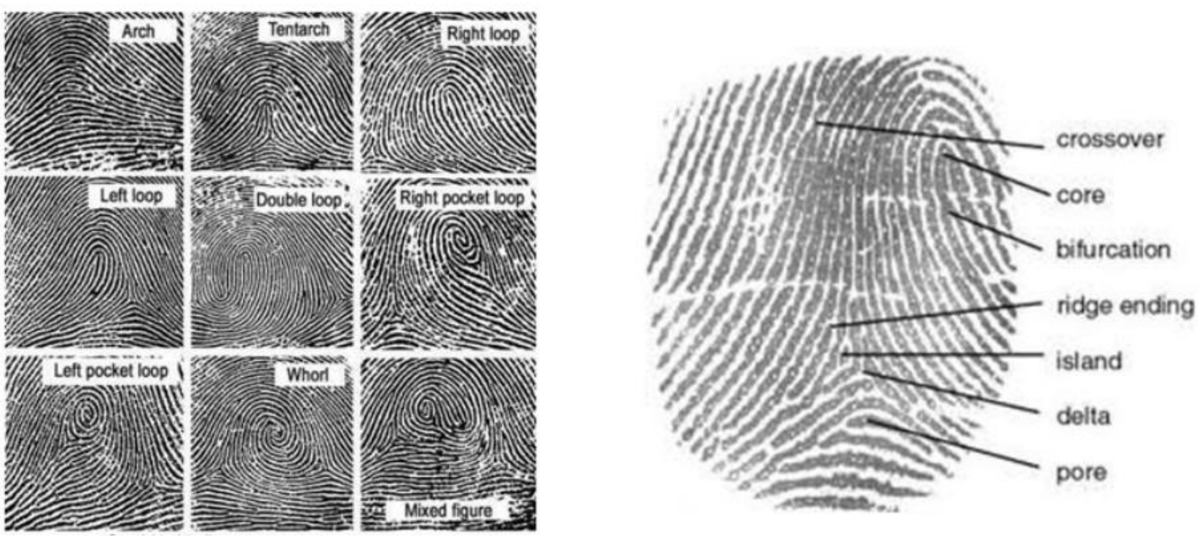

Figure 2.4 Fingerprint Pattern Classes and Minutiae Points (Patel et al., 2013)

2.3 Performance and Evaluation Metrics

There is a difference between overall group performance and individual performance. If performance curves from each sensor in a data collection show the same false acceptance and false rejection rates, this indicates that the same group of people performed identically on each sensor. Some individuals appear to perform better on certain devices or even modalities than they do on others.

A failure to acquire rate, or FTA(R), refers to when a biometric trait is presented but the sensor does not capture an image. This can result from many factors, such as poor not enough applied pressure, the subject not holding still, or the alignment of the biometric with respect to the sensor. Though not always the case, these are often a product of the user's interaction or the environment rather than the technology. A failure to extract, or FTX, occurs when a biometric trait is presented, and an image is capture but quality information cannot be extracted. This occurs when the features of the biometric trait are not clear to the extraction and quality software. This can result from similar factors as an FTA and, like an FTA, is often a product of the user's interaction or environment. A true acceptance rate, or TAR, refers to the rate at which the matching algorithm correctly matches and identifies a genuine user as themselves. A false acceptance rate, or FAR, occurs when an impostor attempt is misidentified as a genuine attempt and incorrectly granted access. A false rejection rate, or FRR, occurs when a genuine user is rejected or failed to be identified as themselves. The threshold of the biometric system determines at what rate these will occur, as explained below.

A receiver operating characteristic curve (ROC curve) is a "plot of the rate of 'false positives' (i.e., impostor attempts accepted) on the x-axis against the corresponding rate of 'true positives'" (ISO / IEC JTC 1 SC 37, 2005, p.6). A detection error trade-off curve (DET curve) is a "modified ROC curve that plots error rates on both axes (false positives on the x-axis and false negatives on the y-axis)" (ISO / IEC JTC 1 SC 37, 2005, p.7). ROC and DET curves are used commonly and extensively for biometric system performance analysis. They "show the relationship between sensitivity (the number of true positives divided by the total number of ground-truth positives) and specificity (true negatives divided by ground-truth negatives)" (Park, Goo, & Jo, 2004). DET curves are used to calculate an EER, which reports when the false match rate (FMR) and false non-match rate (FNMR) are equal. In a verification scenario, the FMR is the rate when the matcher falsely identifies two images as being from the same finger from the same person. The FNMR is the rate at which the matcher falsely identifies two images as being from different fingers or different people.

However, there are other factors that impact performance such as an individual's gender and age, and the quality of the data being captured. These covariates are not considered in ROC and DET curves. The performance of the individual can greatly impact population biometric performance and it is important to evaluate the individual since the population performance curves might not properly illustrate what is occurring within a biometric system. Additional characteristics of biometric performance that are measured to evaluate individual performance include "zoo menagerie", animal characteristics (Dunstone & Yager, 2009). The biometric zoo menagerie characterizes an individual's performance by numerically and graphically classifying performance using descriptors of animals. This enables one to assess a system's and an individual's true performance more easily and readily. This allows for a more accurate or optimized system. Regarding fingerprints, a genuine score is the result of a user's fingerprint matched with another image of the same finger from the same user. An impostor score is the result of a user's fingerprint matched with the image of any finger from another user or a different finger from the same user. For genuine matches, genuine scores should be high and impostor scores should be low.

The zoo menagerie is credited to Doddington, Liggett, Martin, Przybocki, and Reynolds (1998) and has been used with several biometric modalities, including the face, fingerprint, keystroke dynamics, and voice. They originally identified the following animal classifications: sheep, wolves, lambs, and goats. A goat is an individual who is difficult to match. Goats lie below the 2.5 percentile of average match score. Wolves were easy to match as they possessed match scores above the 97.5 percentile. They produce a higher false acceptance rate as they can imitate others well (Doddington et al., 1998). “A lamb is an individual who is particularly easy to imitate and has characteristics similar to others in the dataset. These animals generate scores similar to everyone, which could lead to false accepts. Sheep are those who have higher genuine scores and lower impostor scores, resulting in lower false match rates and low false accepts” (O’Connor, 2013, p. 16-17). Alternative animal classifications have been identified by others. These include chameleons, worms, doves, and phantoms (Yager & Dunstone, 2010).

Within the Yager and Dunstone (2010) classification, “chameleons” possess high genuine and impostor match scores, meaning that they look like themselves and everyone else. Chameleons are “in the top 25% of the genuine distribution and the top 25% of the impostor distribution” (O’Connor, 2013, p. 17). Doves are “the best performing individuals” and are “in both the top 25% of the genuine distribution and the bottom 25% of the impostors” (O’Connor, 2013, p. 17). “Phantoms” have low genuine and impostor scores and do not match well against anyone, including themselves. This means that they are “in the bottom 25% of the genuine and impostor distributions” (O’Connor, 2013, p. 17). “Worms” (the worst performing individuals) have low genuine scores and high impostor scores (Yager, 2007). This means they do not look like themselves but do look like everyone else. Worms reside in the bottom 25% of the genuine matches and in the top 25% of the impostor matches. *Figure 2.5* displays a zoo plot using the Yager and Dunstone zoo methodology. Genuine match scores are plotted on the x-axis and imposter match scores are plotted on the y-axis. Doves (top right quadrant), phantoms (top left quadrant), worms (bottom left quadrant), and chameleons (bottom right quadrant) are all shown in *Figure 2.5*. The middle portion of the zoo plot contains the “Normal” classified individuals. These individuals do not fit within any of the quadrants (percentiles). However, “the zoo philosophy is not well-accepted in the community because it has not been proven significant” (O’Connor, 2013, p. 16).

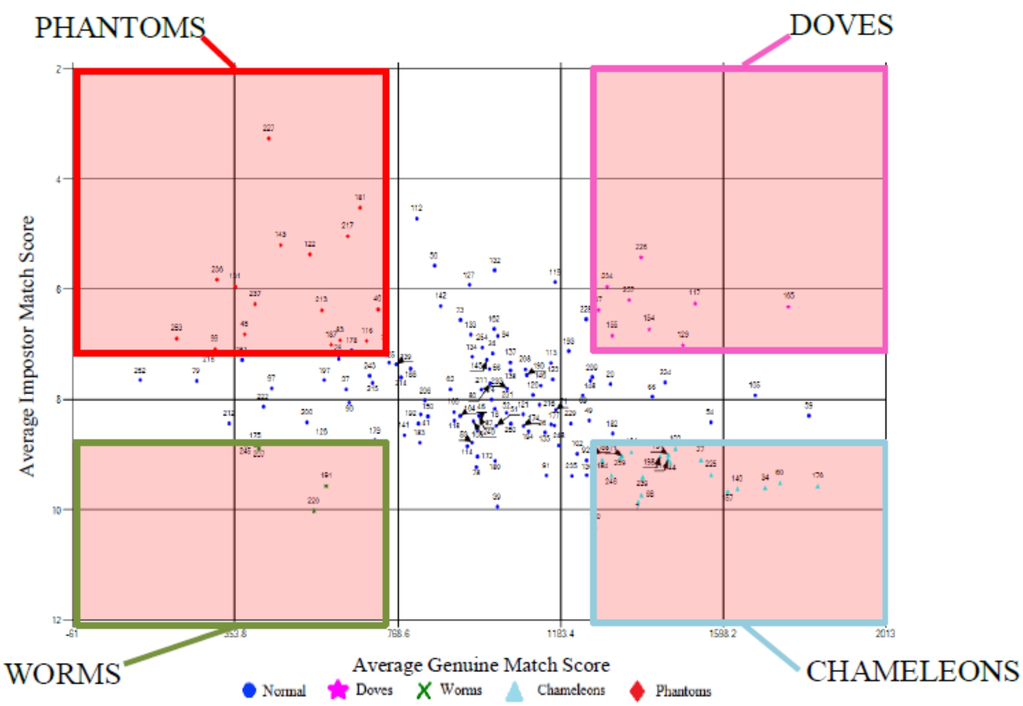

Figure 2.5 Example of Zoo Plot Analysis (O’Connor, 2013)

2.4 Aging

There are various factors that affect the performance of a biometric system’s performance and the individual is just one of them. “Environment, image quality, and device selection play an important role in the successful implementation of a biometric system” (Elliott et al., 2007, p. 5).

Other factors include algorithm, force, time, and aging (O'Connor, 2013). According to Petrelli (2009), user preference and comfort is an additional factor that can impact an individual's biometric performance. There has been an emphasis, until recently, on the sensor and its technology rather than these additional factors. "Research in fingerprinting techniques has favored development of software algorithms over the fingerprint acquisition process to increase recognition system performance" (Petrelli, 2009, p. 3). In biometrics, aging can be separated into two categories, biological and template aging. Biological aging is "the gradual decrease in a system performance caused by the changes suffered by the users' trait in the long-term (which cannot be avoided as is inherent to human nature)" (Lanitis, 2010, p.32) or "the deterioration of the body over time" (Carls, 2009, p. xix). This occurs when the physical biometric trait ages. Individuals experience physical aging with the passage of time, as do their biometric traits and features. Template aging is "when the quality of the match between an enrolled biometric sample and a sample to be verified degrades with increased elapsed time between the two samples" (Fenker & Bowyer, 2011), or "the degree to which biometric data evolves and changes over time, and the process by which templates account for this change" (Carls, 2009, p. xix). This occurs when the template created for future comparisons becomes outdated. As an individual physically ages, the gap in time or technology between the previously stored template and their current physical characteristics can affect the performance and accuracy of the system. In this scenario, the template may need to be updated so it more accurately reflects the current state of the genuine user's biometric trait.

The aging process occurs for various biometric modalities. Some modalities may be considered intrusive, but this perspective may differ between user age groups (Fairhurst et al., 2015). For fingerprint recognition, according to Lanitis (2010), "aging causes reduced skin elasticity that affects the fingerprint scanning process as the contact between dry skins and scanners is not firm" and "at increased age the possibility of observing damaged fingerprints due to wearing and injuries increases." For facial recognition, appearance "is affected considerably by the aging process" and "facial aging is mainly attributed to bone movement and growth and skin-related deformations" (Lanitis, 2010). For iris recognition, the iris is "regarded to be invariant to within-class variation, presenting in that way an ideal biometric feature. The appearance of the iris is formed within a few months from an infant's birth and remains relatively unchanged throughout a person's lifetime. For this reason, the iris is regarded as an aging invariant biometric feature" (Lanitis, 2010). When comparing iris images from sequential and non-sequential visits, Petry (2015) found that while individual scores might change, overall stability did not. There was "statistical stability of the iris within the month duration range" (Petry, 2015, p. 97). However, Bowyer et al. (2007) determined that eye-related diseases such as cataract and glaucoma can affect the stability of the iris over time and impact the accuracy of an iris recognition system, and these diseases tend to increase in prevalence with age.

According to Fairhurst et al. (2015), the effects of biological aging can be categorized into two categories: feature-related and age progression. Feature-related aging issues refer to the differences over time in specific features a biometric system evaluates for a given biometric trait. For example, the ridges in a fingerprint can wear out over time and become less differentiated from valleys. For face recognition, surface texture can change over time in adult faces (such as wrinkles, marks, and scars). The physical shape can change over time in the faces of children. For signature recognition, the velocity, acceleration, and frequency/duration of pen lifts can change with age. Additionally, "template aging occurs as a result of accumulated changes in biometric data during the time which elapses between enrollment (when reference data are first recorded and stored) and authentication (a specific identification/verification event)" (Fairhurst et al., 2015). Previous studies have also been conducted to begin analyzing the effect time and age has on fingerprint recognition at small time intervals.

In Galbally (2019), fingerprint analysis was split into children (0-17 years old), adults (18-25 years old), and elderly (65-98 years old). It was discovered that the elderly group was the most challenging to collect quality data from. For children, the most "problematic group is 0-4. For ages 5-12 fingerprint quality is already acceptable, while for 13-17 it is equivalent to that of adults" (Galbally, 2019, p. 1357). Performance increased drastically between the ages of 0 and 4 years old, increased more slowly between the ages of 5 and 12 years old, and stabilized between 12 and 17 years old. Peak performance appeared to occur for individuals in the adult group (18-25 years old) and the worst performance occurred in the elderly group, particularly between the ages of 81 and 98 years old. It was also found that performance circled back as age increased, as the elderly at 70 years old performed similarly to children between the age of 4 and 5 years old. This was due to fingerprint quality of the elderly at 70 years old being "equivalent to that of 4-5 years old children" (Galbally, 2019, p. 1358). Sickler (2005) also found that the elderly population, ages 55 and older, did not perform well. The increase of age and corresponding decrease in fingerprint moisture "are correlated to lower fingerprint image quality" (Sickler, 2005, p. 1).

Regarding the physical fingerprint, the ridges become brittle with the passage of time which creates a loss of mass due to erosion from friction and air. The moisture in a fingertip also fluctuates over time. While there may not be a lot of mass in the fingertip, it does not take a lot of change to significantly affect the mass with respect to its original value. It has been observed that a decrease in the concentration of lipids, unsaturated fatty acids, triglycerides, cholesterol, and squalene within the fingertip occurs over time, altering the chemical composition of the finger. This could then negatively affect fingerprint performance and biologically impact the aging of the fingerprint over time (Cadd et al, 2015, p. 224-226).

According to Carls et al. (2008), there have been studies done to evaluate the effects of template aging, however, most of these have focused on face recognition. Ryu et al. (2007) conducted an analysis of fingerprint template aging, measuring image quality and matching performance (EER and genuine match score). Ling et al. (2007) conducted a fingerprint recognition study over time, analyzing 86 individuals over a 16-week period. According to Fairhurst et al. (2015), one common obstacle with studies such as these is the lack of datasets with a long enough time lapse. Determining the impact that the template may have on image quality and performance could help decide whether spending the time and money on template renewal is worth it or not. There may be differing levels of force applied across presentations of the same finger from the same user. Matching algorithms are proprietary, so while they each perform the same function, the method by which they do so can vary across algorithms. There are many layers, often more than can be anticipated or predicted.

Template aging is very hard to quantify and evaluate due to numerous covariates over any given period. A robust calculation of this phenomenon is very difficult to obtain due to these numerous covariates and their difficulty to control experimentally (Harvey et al., 2017). Harvey et al. (2017), identified biometric permanence with regards to false non-match ratio (FNMR), since most biometric modalities do not experience large changes in FNMR. Simplifying the evaluation method and first quantifying the effects of time, in general, might be the first step in understanding this effect. Additionally, renewing the template periodically could potentially reduce or alleviate the negative effects produced by template aging (Ross, 2004). There are different sizes of templates and different speeds of template creation. These factors could potentially impact the performance of a biometric system. Regarding fingerprints, the fingerprint can become worn over time. Changes or advancements in technology and scarring could also necessitate the renewal of the template. This could be time consuming and expensive.

2.5 General Infant Biometrics

Previous studies have analyzed the effects of aging on biometric performance, however, there has been difficulty in obtaining a large enough sample size and a great enough time difference. According to Michalski et al. (2018), the National Institute of Standards and Technology (NIST) conducts Facial Recognition Vendor Tests (FRVTs) to evaluate biometric performance. “These evaluations have predominantly focused on images of adults however, studies that have been conducted on images of children have found that algorithm performance is consistently lower than with images of adults. This is likely due to the considerable amount of facial growth occurring in childhood” (Michalski et al., 2018, p. 217). However, it is difficult to evaluate facial recognition with children due to smaller sample sizes. Other variables such as pose, illumination, and expression could also be impacting performance. It is difficult “to determine how age and age variation impacts on performance throughout childhood” and “it is critical to determine how performance changes throughout childhood at the lowest level possible” (Michalski et al., 2018, p. 217).

Fingerprints were the first biometric used to attempt to identify infants and children. “In 1899, Galton first captured ink-on-paper fingerprints of a single child from birth until the age of 4.5 years, manually compared the prints, and concluded that the print of a child at the age of 2.5 years would serve to identify him ever after” (Jain et al., 2016, p. 1). However, according to Bharadwaj et al. (2010), Corby et al. (2006), Jain et al. (2014), Jain et al. (2016), Jain et al.

(2017), Jia et al. (2010), and Jia et al. (2012), clean and precise infant biometrics are extremely difficult to obtain due to an infants' rapidly changing physical growth and development, their natural unintentionally uncooperative behavior, and the fact that most biometric systems are designed with a population of adults in mind.

For footprint recognition, it was found that it was easier to collect images when infants were asleep rather awake (Jia et al., 2012). However, this is not always practical. For palmprint recognition, it was found that ink and paper methods were not effective at all, with the best results coming via an optical finger/palm sensor (Weingaertner et al., 2008). Additionally, the performance was less accurate than fingerprint recognition. For ear recognition, it was found that the performance of samples collected from infants were very similar to the performance of samples collected from adults (Tiwari, Singh, and Singh, 2011). A big challenge with ear recognition is the lack of data on both adults and infants, but especially on adults. This clouds the viability or ability to extrapolate results for an entire population, as ear recognition is not widely used. Since there already are established modalities for collecting adult biometrics and the identification of new modalities or methods focuses more on infants than adults, this mismatch creates an unbalance which makes it difficult to compare infant performance against adult performance. This currently limits our understanding of ear recognition on infants.

For face recognition, neutral facial expressions are typically preferred, and infants exhibit a wide range of facial expressions. They also struggle to comprehend directions and instruction. Movement can cause blurry images and closed eyes can make it difficult for the algorithm to detect the face. It was found that when quality face images were captured the system performed decently well, however it was very difficult to capture high quality face images (Bachenheimer, 2016). Closed eyes obviously negatively impact iris recognition performance as well. However, iris recognition does show promise as possibly the most accurate method of biometric recognition (Corby et al., 2006). It was found that marginal quality images were still able to be successfully matched, and it was rare than an infant had two low-quality and unusable irises. The biggest issue was a high failure to acquire rate, but when successfully captured, it tended to work well compared to other infant biometric modalities.

One of the largest challenges when it comes to infant biometrics is the sheer difficulty in obtaining a large enough sample size to test and design biometric systems. Without a large enough test sample, it is hard to fine tune a system for that specific population (Fairhurst et al. (2015). Additionally, there is the ethical debate of collecting biometric data (Rebera & Mordini, 2013). This, and a lack of motivation from researchers due to the natural non-cooperation of infants, increases the difficulty of infant data collection.

2.6 Infant Fingerprint Recognition

In infants, the quality of captured fingerprint images tends to be very poor (Jain et al., 2016). Since palmprint and footprint recognition both use ridge-based biometric feature extraction like fingerprint recognition, many of the same challenges apply to these modalities as well. Though, according to Jain et al. (2017), a high-resolution fingerprint scanner made for the specific purpose of collecting infant fingerprints may mitigate some of these issues.

In Jain et al. (2014), fingerprint images were collected in a controlled environment at Michigan State University, using the Digital Persona U.are.U 4500 (same device that will be used in this study). Fingerprints were collected from infants 0-4 years old, over five visits. The first and last visits for each infant were one week apart. Images of each infant’s right and left index fingers and thumbs were collected. A latent algorithm and live-scan algorithm were used to match these images and it was discovered that the latent algorithm performed better than the live-scan algorithm. This was due to infant fingerprints sharing quality characteristics with latent adult fingerprints and suffering from the same issues as latent adult fingerprints, such as incomplete prints or poor ridge and valley contrast.

Matching algorithms and devices have been primarily developed for the adult population and were tested on the adult population. This results in unanticipated issues, or possibly ones that cannot be addressed due to the system's operational baseline as determined by the 'typical' adult fingerprint. Another cause of this discrepancy could be the general difficulty of successfully matching or extracting quality metrics from latent fingerprints. Since latent prints are dealing with whatever portion of a fingerprint that was been left behind, there are many uncontrollable variables. The algorithm, even when accounting for this, could have difficulty in reading information with any degree of certainty, for both infant and adult latent fingerprints.

In Jain et al. (2016), fingerprint images were collected from 66 infants between 0-6 months old. Three images each of both left and right thumbs were taken from each infant that participated in this study. Images were collected over two separate visits. A set of images was obtained on the infant's first visit and another set was obtained between 2 and 4 days later. For analysis in a verification scenario, the first set of images for each infant was used as templates and the images obtained in the infant's subsequent visit were used as the current presentation of the same fingers to verify they are who they are supposed to be. The frequency of true acceptances and false acceptances (equivalent to a false match rate) were evaluated. For the identification scenario, each fingerprint was compared to a background database of fingerprints belonging to known subjects. Encouraging results were discovered with their custom fingerprint sensor. This sensor was designed to be compact and high-resolution to detect the details of an infant fingerprint more successfully and accurately. The dimensions of the sensor were 7 cm x 3 cm x 7.5mm, and the resolution was 1,270 ppi. This was much higher than the ~500 ppi resolution of standard commercially available fingerprint sensors that have been used in the other mentioned studies. Using this custom sensor, infants older than 4 weeks old produced a true acceptance rate (TAR) of 83.55% and a false acceptance rate (FAR) of 1%, in a verification scenario. In an identification scenario, an 79.95% success rate was achieved. Though with infants younger or equal to 4 weeks old, the results were still unfavorable. Here, the TAR was only 54.55% in a verification scenario and the success rate in an identification scenario was only 44.05%. However, to the researchers' knowledge this was the first time that fingerprints with visible ridge details were captured from infants as young as 6 hours old.

Jain et al. (2017) used the same custom fingerprint sensor that was used previously in Jain et al. (2016) and built upon that research. In Jain et al. (2017), fingerprint data was collected from 309 subjects aged 0-5 years old. These children were separated into three different age groups: 0-6 months old, 7-12 months old, and >12 months old. Three data sets were used, all consisting primarily of fingerprint images captured from infants 0-6 months old. Dataset A (204 infants) consisted of fingerprint images collected on both the custom 1,270 ppi sensor and a standard 500 ppi sensor, over four visits. Dataset B (65 infants) consisted of fingerprint images collected on the custom sensor only, over three visits. Dataset C (40 infants) consisted of fingerprint images collected on the custom sensor only, over two visits.

It was discovered that children fingerprints do "possess the salient features necessary to uniquely recognize each child" (Jain et al., 2017, p.12). Additionally, it was determined that it is "possible to capture a child's fingerprints with sufficient fidelity for recognition" at the age of 12 months old (TAR = 99.5% at FAR = 0.1%) when using a standard sensor with a 500 ppi resolution, and at 6 months old (TAR = 98.9% at FAR = 0.1%) when using the custom sensor with a 1,270 ppi resolution (Jain et al., 2017, p.12). A standard resolution sensor would be adequate for children 12 months and older. However, for fingerprint recognition on children as young as 6 months old, the higher resolution custom sensor would be necessary. It was also found that a child's age at the time of enrolment played a larger role in affecting genuine match scores than the time lapse between enrolment and re-presentation of the finger, and the genuine match scores did not significantly decrease over the time lapses of 6-12 months. Matching was more affected by the physical age of the child than the passage of time. This was likely due to younger children supplying poorer quality enrolment images which later affected the matching performance. However, once quality information was captured and stored as a template, the matching performance was largely unaffected (Jain et al., 2017).

Figure 2.6 shows images of the left thumbprint belonging to the same child taken at three separate times. From left to right, the infant's thumbprint was captured at 1 day old, 3 months old, and 6 months old. It is hard to differentiate between the ridges and valleys for each fingerprint.

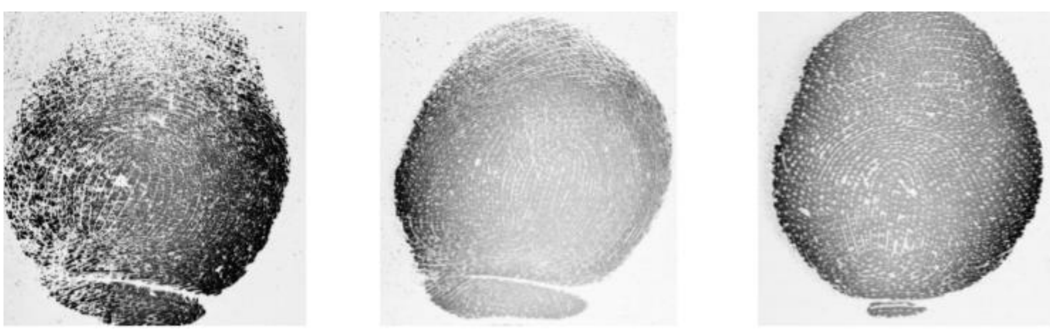

Figure 2.6 Example of Left Thumb Images from the Same Infant Child (Jain et al., 2017)

Figure 2.7 shows a teenager's fingerprint. The ridges and valleys are much more visible and distinguished.

Figure 2.7 Example of Quality of a Teenager's Fingerprint (Jain et al., 2017)

Figure 2.8 shows an example of a fingerprint captured from the elderly. The ridges and valleys are visibly interfered with by scars and creases, from aging and wear and tear on the fingers.

Figure 2.8 Example of a Fingerprint Capture from the Elderly (Jain et al., 2017)

2.7 Challenges with Collecting Infant Data

It is difficult to collect biometric data from infants (Jain et al., 2014). This is mostly due to the difficulty of getting infants to properly interact with a fingerprint sensor. Infants' fingers get wet from them putting their fingers in their mouth, they often time clench and unclench their fists and move their hands and fingers in seemingly random patterns, can be balled up (Jain et al., 2014). Additionally, the spacing of ridges and valleys in infants' fingerprints is smaller and more compact than that of an adult fingerprint. This creates an additional step in feature extraction, accounting for this spacing discrepancy (Jain et al., 2014).

According to the Dutch Ministry of the Interior and Kingdom Relations (2005), infants younger than 8 or 9 months old often make strong fists which are difficult to open and inhibit the fingerprint data collection process. Children's hands are often moist from them sucking their thumbs and need to be dried off often, however drying off their hands does not fully address the issue of the skin becoming softer. This resulting malleable nature of the skin means a much higher difficulty in capturing a good quality fingerprint image. It was difficult, but possible, to capture fingerprints from children around 3 or 4 years old. The primary finger that was successfully capture was the thumb. This was presumably because it has a larger surface area than the other fingers, which results in more potential data points that can be captured and extracted. It was exceedingly difficult to collect fingerprint data from children younger than 3 or 4 years old.

It was also said to be "virtually impossible to obtain fingerprints from children aged under 4 years" (Dutch Ministry of the Interior and Kingdom Relations, 2005, p.25). This conclusion was derived from the failure to capture rate exceeding 50% until children reached 4 years of age. And even at 50%, that is a very high failure to capture rate and a very low successful capture rate.

CHAPTER 3. METHODOLOGY

The methodology was divided into four sections, an introduction, the data collection procedures, device specifications, and the data calculation and analysis methods.

3.1 Introduction

Fingerprint recognition is commonly viewed as unviable for use with the infant population (Dutch Ministry of the Interior and Kingdom Relations, 2005). The research question being addressed is as follows: is there a difference in infant fingerprint performance and image quality metrics, between different age groups (0-6 months, 7-12 months, and >12 months old), using the same optical sensor? Additionally, is there stability in individual infant fingerprint performance across age groups for children who aged into a separate age group during the longitudinal data collection? This research is quantitative and will analyze each age group as a separate population.

3.2 Infant Data Collection

The data for the three age groups (0-6, 7-12, and >12 months old) used in this secondary analysis was captured in multiple visits as part of a longitudinal multimodal infant data collection. Hutchison (2018) evaluated infant iris recognition using data collected from the same longitudinal study. The iris analysis was separated into three infant age groups: “0 to 6 months old, 7 to 12 months old, 13 to 24 months old” (Hutchison, 2018, p. 55), and thus, for comparison to this analysis, the same age groups were chosen

Subjects were separated into these different, independent age groups based upon their current age at the time of data collection. Some individuals appeared in one age group at the beginning of the longitudinal data collection and appeared in another age group by the end. This created an element of overlap between age groups when a subject aged out of one group and entered into another. This overlap was evaluated to determine whether individual stability was present across age groups. The data collection procedures from the original study that this research is utilizing are described below.

The parents/guardians were shown how to place the finger on the device. For children under the age of 24-months-old, the parent/guardian would hold the fingerprint sensor up to the child with one hand, and then with their other hand, place the infants right or left index finger onto the device. The number of attempts were not defined due to many subjects needing numerous attempts in a short period of time before being upset. This was determined by parent/guardian or test administrator’s level of comfort. Images of both the left and right index fingers were captured for each child, and the parent/guardian decided which index finger to start with since the child would sometimes have a toy in one of their hands. Left and right index fingers for each child were treated independently since the left index is a different finger from the right index. This data collection consisted of 2,769 total images, collected from 38 children in the 0-6 months old group, 44 children in the 7-12 months old group, and 58 children in the >12 months old group. *Table 3.1* illustrates the breakdown of these age groups. The expected numbers column accounts for when and where an FTX occurred.

Table 3.1 Separation of Subjects by Age Group

Age (Months)	Total Subjects	Total Images	Expected Images
0-6	38	523	574
7-12	44	918	926
>12	58	1,269	1,269

3.3 Device Specifications

The device used for all fingerprint captures was the Digital Persona U.are.U 4500, as shown below in *Figure 3.1*. The specifications for this sensor are listed below in *Table 3.2*.

Table 3.2 Digital Persona U.are.U 4500 Specifications

Specification	Value
Manufacturer	DigitalPersona, Inc.
Sensor Type	Optical
Interaction	Touch
Platen Size (mm)	15 x 18
Resolution (ppi)	512
Illumination	Blue LEDs
Device Size (mm)	65 x 36 x 16
Connection	USB 2.0
Supported OS	Windows, Linux, Android

Figure 3.1 Image of Digital Persona U.are.U 4500

3.4 Calculation and Analysis Method

Image quality and minutiae count were the evaluated image quality metrics. Image quality scores were computed by the Neurotechnology 10 SDK, on a scale from 0 to 100. Images that fail to produce quality information were categorized as an FTX. An FTX is derived from images that failed to produce quality data divided by the total number of images. Genuine and impostor match scores were the evaluated performance metrics. This was demonstrated using DET curves and zoo analysis. A one-way MANOVA was used to measure statistical significance in different image quality metrics, by age group. Additionally, a one-way MANOVA was used to measure statistical significance in genuine and impostor match scores by each age group. A one-way MANOVA was selected since both analyses include one independent variable with 3 levels (age group) and more than one dependent variable (2). For both analyses, the independent variable was age group (0-6, 7-12, and >12 months old). For the analysis of the image quality metrics, the dependent variables were image quality and minutiae count. For the MANOVA analysis of the matching performance, the dependent variables were genuine and impostor match scores.

The zoo menagerie and stability score index (SSI) were used to evaluate stability, for children who overlapped between more than one age group during the multimodal data collection. The zoo menagerie was used to graphically evaluate stability and the SSI was used to numerically evaluate stability. The SSI formula, shown below in *Figure 3.2*, evaluates stability by providing a numeric value between 0 and 1, where 0 indicates no difference is present and 1 indicates the maximum difference possible occurred.

$$\boldsymbol{S.S.I_i} = \frac{\sqrt{\left(x_{i_2} - x_{i_1}\right)^2 + \left(y_{i_2} - y_{i_1}\right)^2}}{\sqrt{(x_{max} - x_{min})^2 + (y_{max} - y_{min})^2}}$$

Figure 3.2 Stability Score Index Formula (O'Connor, 2013)

The SSI calculates the individual stability between two levels of any one variable. In this study the variable was age and the levels were the specific age group. "X_1 and X_2 represent the genuine match scores for the two levels examined. Y_1 and Y_2 represent the individual's impostor match scores from each level. X_{max} and X_{min} represent the maximum obtained genuine score and minimum possible score that was seen at all levels. Y_{max} and Y_{min} represent the maximum obtained impostor score and the minimum possible score that was seen at all levels. The numerator represents the individual's movement over the two levels and the denominator represents the maximum possible movement amongst all levels" (O'Connor, 2013, p. 52). The result of this calculation is an index value between 0 and 1, where 0 indicates most stability and 1 indicates the most instability.

Again, the research questions were as follows:

1. Is there a difference in fingerprint matching performance between infant age groups?
   Metrics include: FTX rates, Genuine and Impostor match scores, DET curves and EER, Zoo analysis
2. Is there a difference in fingerprint image quality metrics between infant age groups?
   Metrics include: Image Quality and Minutiae Count
3. Is there individual stability in performance for infants who aged into a different age group during the longitudinal data collection?
   Metrics include: Zoo Analysis, Stability Score Index

And the hypotheses were as follows:

1. There is no difference in infant fingerprint performance between the three age groups.
2. There is no difference in infant fingerprint image quality metrics between the three age groups.

3. There is stability in individual infant fingerprint performance for those who aged into a different group during the longitudinal data collection.

3.5 Methodology Summary

This study was performed to answer the previously stated research question, “is there a difference in fingerprint matching performance between different age groups (0-6 months, 7-12 months, and >12 months old), using the same optical sensor?”, by comparing genuine and impostor match scores from children in three different age groups. Whether there was stability in individual infant fingerprint performance across age groups for children who aged into a separate age group during the longitudinal data collection was also analyzed. Using the same optical sensor isolates age group as the variable of interest. Measuring and quantifying changes and differences in infant fingerprint performance, by age group, would be very impactful. This would add to the body of knowledge, could discover differences that may point toward specific age groups to use or avoid for infant fingerprint recognition, and could assist in the understanding of biometric systems and their successful implementation and integration with the infant population.

CHAPTER 4. RESULTS

The analysis was divided into four sections, the data cleaning procedures, analysis of the matching performance (DETs), analysis of the image quality metrics (one-way MANOVA), and analysis of genuine and impostor match scores (one-way MANOVA and Zoo Plots).

4.1 Data Cleaning Procedures

The data used for this secondary analysis consisted of 2,769 total images. Fingerprint image quality metrics were calculated using the Neurotechnology 10 SDK. All fingerprint images that failed to produce quality data were labeled and removed from the analysis. These images were unusable for analysis since they could not even produce a score zero. For inclusion in the secondary analysis, subjects needed to have at least two images that were captured and were able to produce quality data. The failure to extract rates for each age group are shown in *Table 4.1*. The remaining images with valid quality data were used to create templates, which were input to the MegaMatcher 10 matching algorithm to assess performance.

Sample images from this dataset are displayed below in *Figure 4.1*, illustrating examples of images that were removed from the analysis due to a failure to extract. This meant that the image was captured but was of such poor quality that the necessary fingerprint features used to compute quality could not be extracted. For the data collection, the quality software threshold for acquiring a minimum number of minutiae points was set to zero so that a fingerprint was captured. This means that there were likely less attempts resulting in a failure to acquire (FTA) than there would have been otherwise, however there were also likely more FTXs than there would if the minutiae count threshold had been turned on. Some common quality issues with fingerprint images were found with these images in *Figure 4.1*. Images appear to be smudged or the valleys appear to be missing. This is likely due to there being too much or uneven pressure being applied during the fingerprint capture process. Another issue that is evident here includes the angle or placement location of the finger during the capture process.

Figure 4.1 Failure to Extract (Left Index from Subject 27, Right Index from Subject 42)

This left 2,710 total images, of which there were 254 left index and 269 right index images from children in the 0-6 months group, 458 left index and 460 right index images from children in the 7-12 months group, and 622 left index and 647 right index images from children in the >12 months group. There were 60 subjects that overlapped between the 0-6 and 7-12 months age groups, 62 subjects that overlapped between the 7-12 and >12 months age groups, and 40 subjects that overlapped between all three age groups.

4.1.1 Failure to Extract Rate by Age Group

The failure to extract rates for each age group decreased as age increased, as shown in *Table 4.1*.

Table 4.1 Failure to Extract Rate by Age Group

Age (Months)	Failure to Extracts	Total Images	FTX Rate (%)
0-6	51	574	8.89
7-12	8	926	0.86
>12	0	1,269	0.00

4.2 Analysis of Image Quality Metrics

Sample images from this dataset are displayed below in *Figures 4.2-4.6*. There are examples of fingers, with poor, fair, good, very good, and excellent quality, as computed by the Neurotechnology 10 SDK. The two examples of excellent quality, *Figure 4.6,* show more even applied pressure, a more central finger placement on the sensor, and a more complete image with proper contrast of the ridges and valleys.

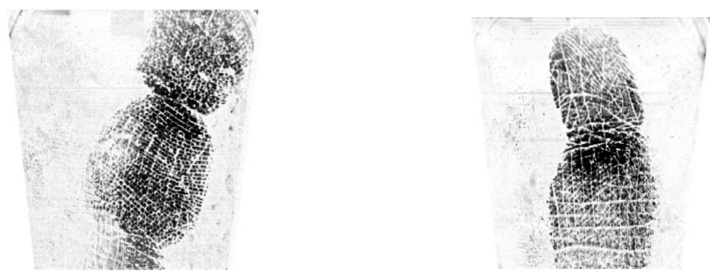

Figure 4.2 Poor Quality (Left Index from Subject 5, Right Index from Subject 7)

Figure 4.3 Fair Quality (Right Index from Subject 64)

Figure 4.4 Good Quality (Left Index from Subject 71, Right Index from Subject 12)

Figure 4.5 Very Good Quality (Left Index from Subject 31, Right Index from Subject 12)

Figure 4.6 Excellent Quality (Left Index from Subject 75, Right Index from Subject 89)

To analyze whether there is a difference in image quality and minutiae count between age groups (0-6, 7-12, and >12 months), a one-way MANOVA was run at a significance level of .05. The independent variable was age and the dependent variables were image quality and minutiae count. Since the one-way MANOVA only indicates if a difference in group means exists and not which specific group means are different, a Tukey (equal variances are assumed) or Games-Howell (equal variances are not assumed) post hoc test was run in order to identify which specific group means are significantly different.

Assumptions for running a one-way MANOVA were met. There were two or more continuous dependent variables and the independent variable consisted of two or more categorical and independent groups. Additionally, there was independence of observations, an adequate sample size (more cases in each group than the number of dependent variables being analyzed), and there was no multicollinearity.

The Levene's test for equality of variances was run to determine homogeneity of variance. The Levene's test was significant for image quality ($p < .001$), meaning there was a statistical difference in variances for image quality. Thus, the Games-Howell test was used for post hoc analysis. The Levene's test was not significant for minutiae count ($p = .14$), meaning there was not a statistical difference in variances for minutiae count. Thus, the Tukey test was used for post hoc analysis. Descriptive statistics for the image quality analysis are displayed below in *Table 4.2*. Image quality and minutiae count appeared to increase with age. Additionally, since image quality scores were computed on a scale from 0 to 100, the quality was still poor despite the large increase with age. The mean image quality peaked with the >12 months age group, but only produced an average quality score of 42.51 out of 100.

The one-way MANOVA found that there was a statistically significant difference in quality based upon a child's age, F (4, 5412) = 306.20, $p < .001$; Wilk's $\Lambda = .67$, partial $\eta^2 = .19$. Additionally, age group had a statistically significant effect on both image quality (F (2, 2707) = 661.47; $p < .001$; partial $\eta^2 = .33$) and minutiae count (F (2, 2707) = 152.88; $p < .001$; partial $\eta^2 = .10$). The Games-Howell test showed that image quality was significantly different between all age group comparisons ($p < .001$) and the Tukey test showed that minutiae count was significantly different between all age group comparisons ($p < .001$).

Table 4.2 Descriptive Statistics Summary: Mean Image Quality

Metric	Age (Months)	Mean
Quality	0-6	17.83
	7-12	24.25
	>12	42.51
Minutiae	0-6	28.78
	7-12	31.21
	>12	35.64

4.3 Analysis of Matching Performance (DETs)

DET curves and EERs were computed using Oxford Wave Research Bio-Metrics OWR performance software. These are displayed below in *Figures 4.7-4.8*. *Figure 4.7* displays the DET curve and EER for all children in this study. This served as a baseline for comparing the performance of each individual age group.

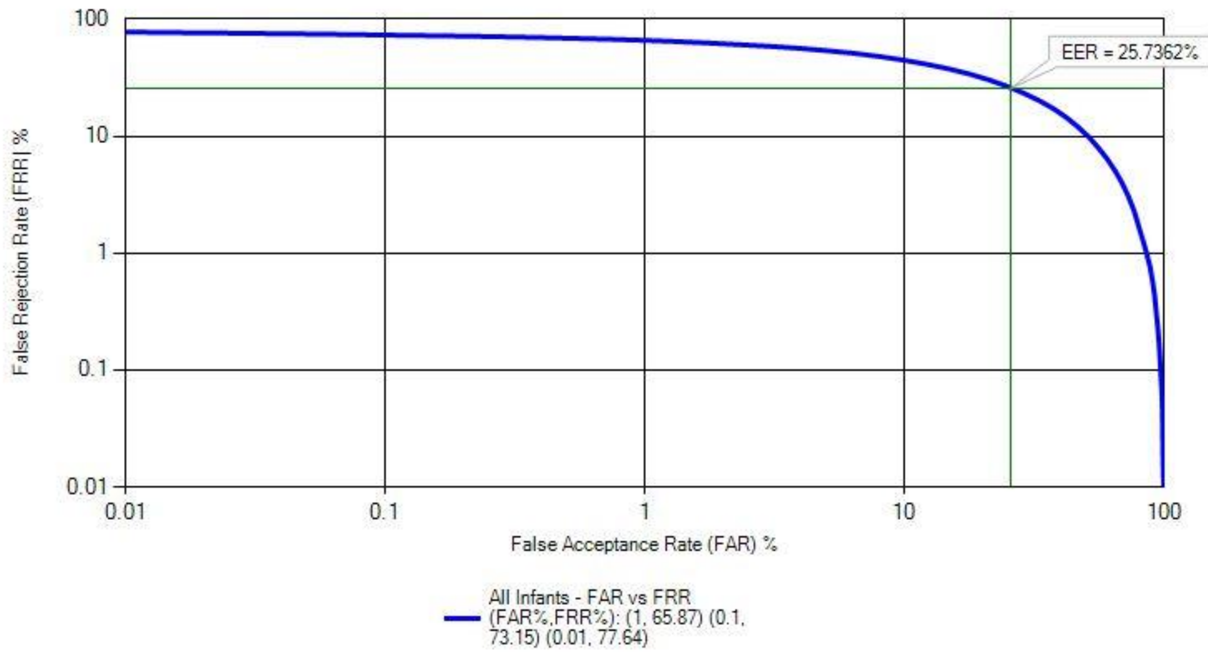

Figure 4.7 Overall DET Curves for All Subjects

Figure 4.8 displays the DET curves and EERs for each age group, 0-6 months, 7-12 months, and >12 months. The EERs by age group are displayed in *Table 4.3* and are compared to the overall EER which included all children in *Table 4.4*.

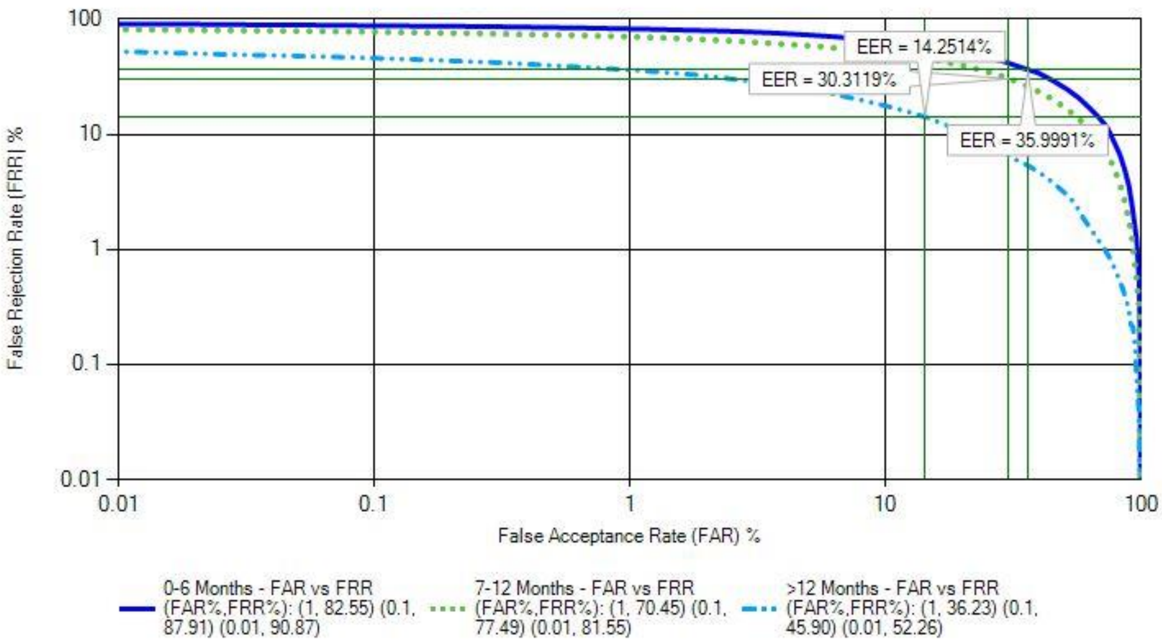

Figure 4.8 DET Curves by Age Group

Table 4.3 EERs by Age Group

Age (Months)	EER (%)	FAR (%)	FRR (%)
0-6	36.00	0.1	87.91
7-12	30.31	0.1	77.49
>12	14.25	0.1	45.90

Table 4.4 EERs: Each Age Group Compared to All

Age (Months)	EER (%)	FAR (%)	FRR (%)
0-6	36.00	0.1	87.91
7-12	30.31	0.1	77.49
>12	14.25	0.1	45.90
All Groups	25.74	0.1	73.15

There was a difference in equal error rate between age groups. The overall system error decreased as the age of the children increased. The largest change occurred between the 7-12 months and >12 months groups (30.31% to 14.25%), whereas the difference between the 0-6 months and 7-12 months group was much smaller (36.00% to 30.31%). Additionally, as shown in *Table 4.4*, children over the age of 12 months were the only age group to perform better than the overall performance with all children included for both the EER and FRR metrics. The other two age groups had higher EERs and FRRs at the same FAR. Therefore, in accordance with Jain et al. (2017), 12 months old appeared to be a benchmark age for using fingerprint recognition on children.

4.4 Analysis of Genuine and Impostor Match Scores

Statistical analysis of the genuine and impostor match scores was conducted by running a one-way MANOVA, since there was one independent variable (age group) and two dependent variables (genuine match score and impostor match score).

4.4.1 Differences in Genuine and Impostor Match Scores Across the Groups

To analyze whether there is a difference in genuine and impostor match scores between age groups (0-6, 7-12, and >12 months), a one-way MANOVA was run at a significance level of .05. Since the one-way MANOVA only indicates if a difference in group means exists and not which specific group means are different, a Games-Howell (equal variances are not assumed) post hoc test was run in order to identify which specific group means are significantly different.

Assumptions for running a one-way MANOVA were met. There were two or more continuous dependent variables and the independent variable consisted of two or more categorical and independent groups. Additionally, all observations were independent, there was a large enough sample (there were more observations than the number of dependent variables), and there was no multicollinearity.

The Levene's test for equality of variances was run to determine homogeneity of variance. The Levene's test was significant for image quality ($p < .001$), meaning there was a statistical difference in variances for image quality. Descriptive statistics for the image quality analysis are displayed below in *Table 4.5*. Image quality and minutiae count appear to increase with age.

The one-way MANOVA found that there was a statistically significant difference in quality based upon a child's age, F (4, 544) = 151.45, $p < .001$; Wilk's $\Lambda = .22$, partial $\eta^2 = .53$. Additionally, age group had a statistically significant effect on both image quality (F (2, 273) = 37.58; $p < .001$; partial $\eta^2 = .22$) and minutiae count (F (2, 273) = 391.14; $p < .001$; partial $\eta^2 = .74$). The Games-Howell test showed that genuine match score was significantly different between all age group comparisons. However, the difference between the 0-6 months and 7-12 months groups exhibited only a minor significance ($p = .05$), whereas all other comparisons exhibited significance ($p < .001$). The Games-Howell test also showed that impostor match score was significantly different between all age group comparisons ($p < .001$).

Table 4.5 Descriptive Statistics Summary: Mean Match Scores

Match Score	Age (Months)	Mean
Genuine	0-6	50.76
	7-12	62.38
	>12	123.96
Impostor	0-6	37.97
	7-12	34.93
	>12	25.53

4.4.2 Zoo Analysis and Stability Score Index

The zoo analysis was conducted to evaluate stability across age groups for children who aged into a separate group during the longitudinal data collection. Individual stability was evaluated since the DET curves showed a difference between age groups, but DET curves do not show whether this difference was caused by the system or the individual. Subjects were separated into different, independent age groups based upon their current age at the time of data collection. Since some individuals appeared in one age group at the beginning of the longitudinal data collection and appeared in another age group by the end, there was some overlap between age groups when a subject aged out of one group and into another. *Table A.1*, attached at the end as Appendix A, displays which subjects aged into a different age bracket and which age groups they overlapped.

A zoo plot shows the performance of each individual in a dataset, relative to all other individuals within the same dataset. Typically, the individuals would be spread out with a higher concentration towards the center of the plot, where subjects with the "normal" animal classification are located. However, for this study, there was a lot of clustering in the top left corner of the plots. Less spread across the individuals represents more consistency. However, this study found that children, especially those in the 0-6 and 7-12 month groups, consistently performed poorly. The tendency for younger children to cluster in the top left corner, classified as "phantoms", meant that they matched poorly against themselves and everyone else in the dataset. This was likely due to very low genuine match scores and the fact that the genuine and impostor scores were very similar, making it difficult for the algorithm to differentiate between the children. *Figure 4.9* illustrates an example of a typical adult zoo plot.

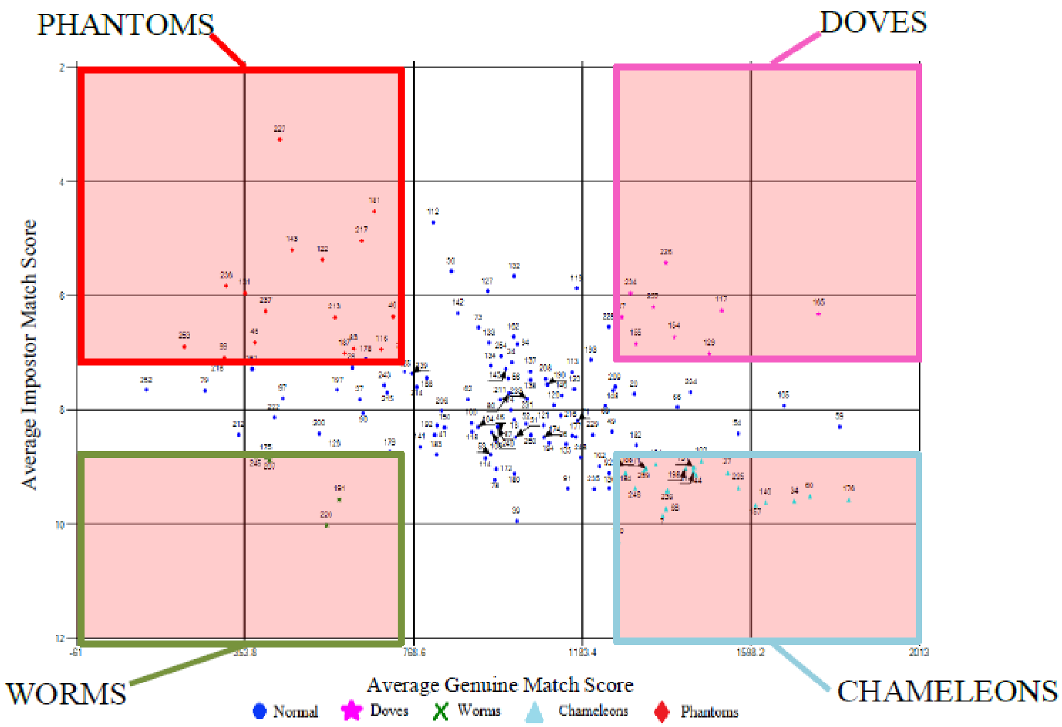

Figure 4.9 Example of Adult Zoo Plot (O'Connor, 2013)

4.4.2.1 <u>Overlap Between 0-6 Months and 7-12 Months</u>

Figures 4.10-4.15 display the zoo analysis for individuals who were in both the 0-6 months and the 7-12 months age groups. *Figure 4.10* shows the zoo analysis on overlapping subjects in the 0-6 months old group and *Figure 4.11* shows the zoo analysis on overlapping subjects in the 7-12 months old group. *Table 4.6* displays the overall animal classification breakdown by age group for subjects who were in both the 0-6 months and 7-12 months groups. *Appendix B* displays the individual breakdown by age of animal classifications and genuine and impostor match scores for each subject.

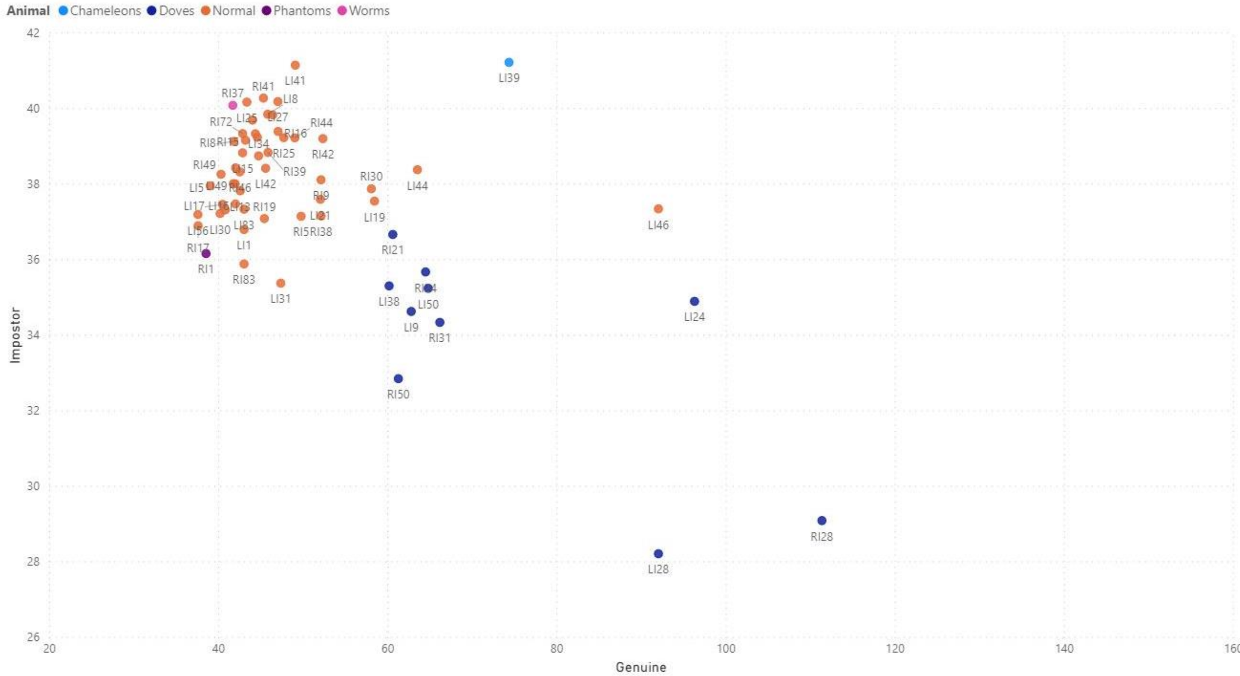

Figure 4.10 Zoo Plot: 0-6 Months Old (0-6 Months and 7-12 Months Overlap)

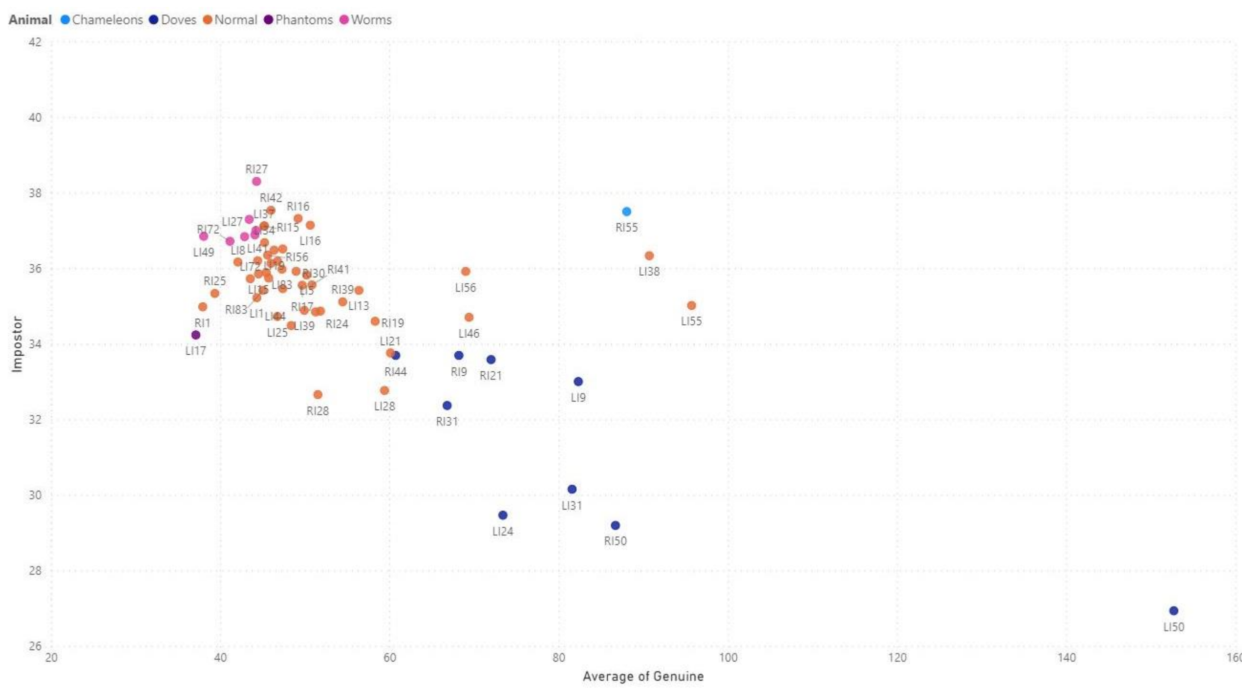

Figure 4.11 Zoo Plot: 7-12 Months Old (0-6 Months and 7-12 Months Overlap)

Table 4.6 Animal Classification Breakdown: 0-6 and 7-12 Months Overlap

Animal Classification	0-6 months	7-12 months
Chameleon	1	1
Dove	10	9
Normal	47	42
Phantom	1	1
Worm	1	7
Total	60	60

Overall, subjects were unstable across the 0-6 months and 7-12 months age groups. Individuals did not always remain the same animal across age groups. For example, subject 39 was a chameleon in the 0-6 months group and a normal classification in the 7-12 months group, where subject 55 was a normal classification in the 0-6 months group and a chameleon classification in the 7-12 months group. Subject 55's genuine match score increased by 41.62 and their impostor match score decreased by 2.32. Subject 55's SSI value was 0.096, much greater than the mean and median SSI values of 0.031 and 0.018, respectively. This illustrated individual instability. Subject 1 was a phantom in the 0-6 months group and a normal classification in the 7-12 months group, where subject 17 was normal classification in the 0-6 months group and a phantom classification in the 7-12 months group. Subject 17's genuine match score decreased by 3.41 and their impostor match score decreased by 3.22. However, this would be considered a borderline case since subject 17's SSI value was 0.011, less than the mean and median SSI values of 0.031 and 0.018, respectively. This demonstrates the importance of not relying solely on match scores and instead using the stability score index to further evaluate individual performance.

It was interesting to note that there was a decrease in normal and dove classifications and an increase in worm classifications, as subjects aged from the 0-6 months group into the 7-12 months group. The genuine and impostor match scores remained similar between the 0-6 and 7-

12-month groups. The increase in worm classifications was likely a result of their lack of improvement as other subjects improved. This means that while the DET curves showed a slight improvement in EER between 0-6 months and 7-12 months, these individuals now classified as worms performed worse compared to others in the 7-12 months group than they did in the 0-6 months old group, despite similar match scores.

4.4.2.1.1 Stability Between 0-6 Months and 7-12 Months

Figures 4.12 and *4.13* illustrate an example of individual stability across age the 0-6 months and 7-12 months age group.

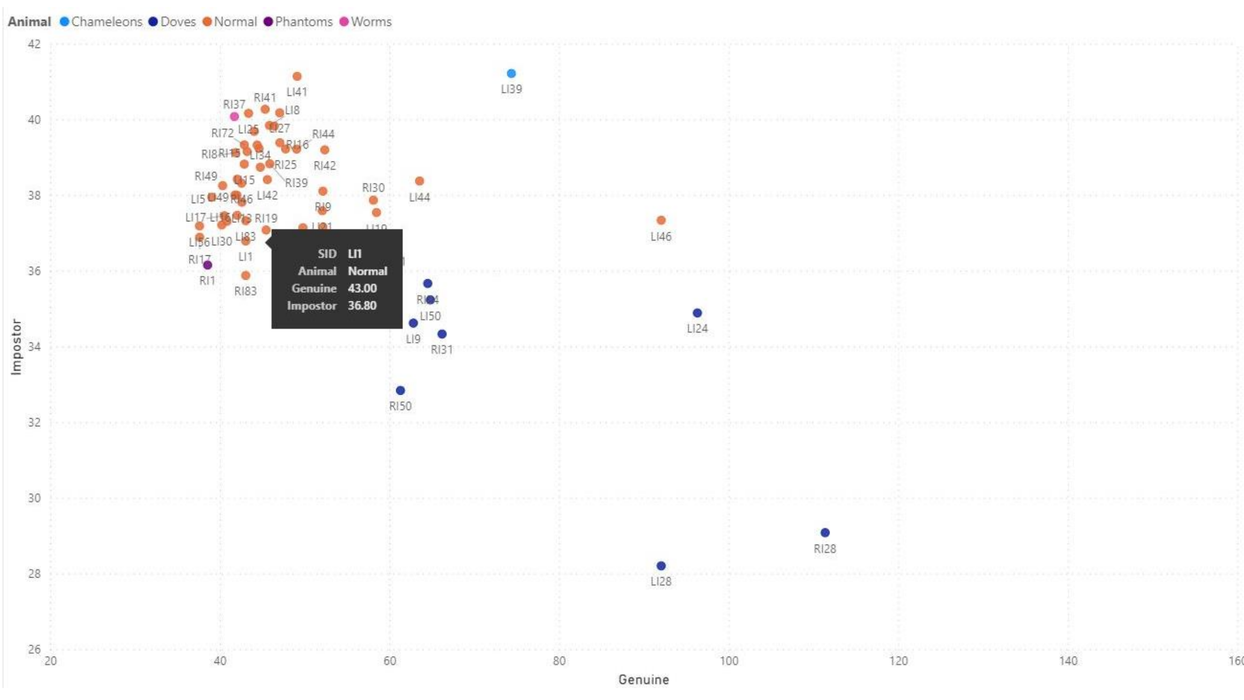

Figure 4.12 Zoo Plot: Subject LI1 at 0-6 Months Old

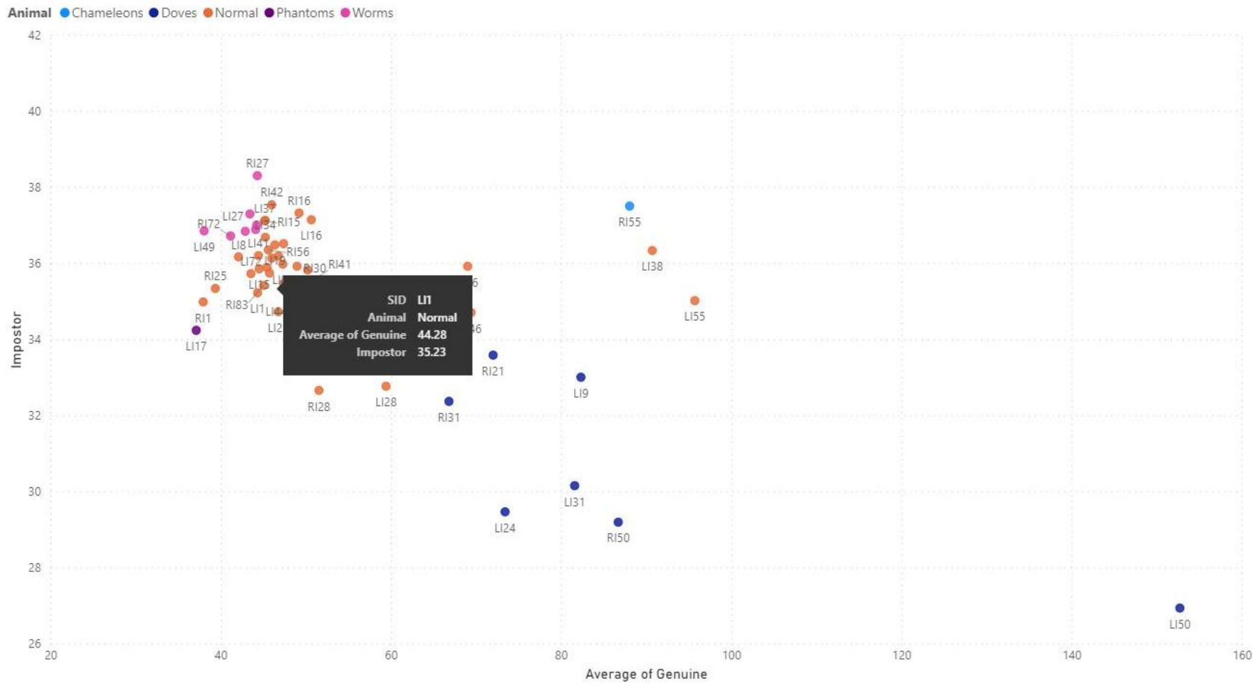

Figure 4.13 Zoo Plot: Subject LI1 at 7-12 Months Old

In *Figure 4.12*, while in the 0-6 months age group, the left index finger of subject 1 had an average genuine match score of 43.00 and an average impostor match score of 36.80. In *Figure 4.13*, while in the 7-12 months age group, the left index finger of subject 1 had an average genuine match score of 44.28 and an average impostor match score of 35.23. The minimal change in match scores across age groups resulted in an SSI value of 0.005. In comparison, the mean and median SSI for the 0-6 months and 7-12 months overlap was 0.031 and 0.018, respectively. This represented stability in subject 1's individual fingerprint performance.

4.4.2.1.2 Instability Between 0-6 Months and 7-12 Months

Figures 4.14 and *4.15* illustrate an example of individual instability across age the 0-6 months and 7-12 months age group.

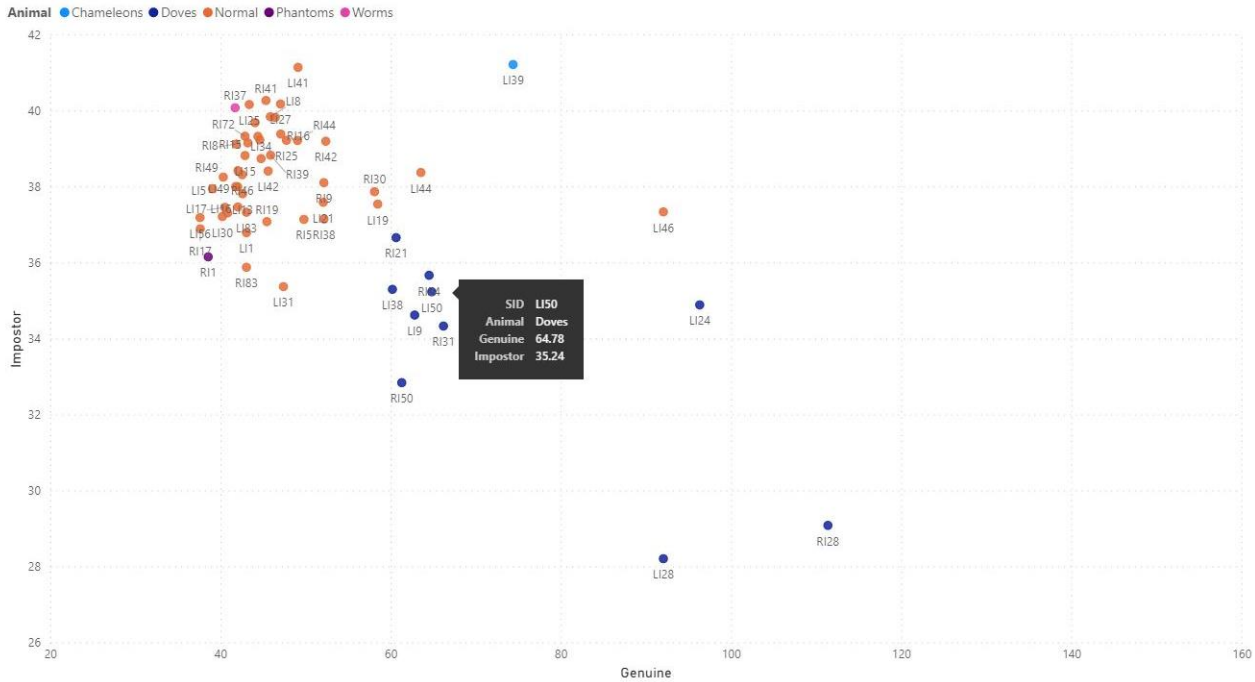

Figure 4.14 Zoo Plot: Subject LI50 at 0-6 Months Old

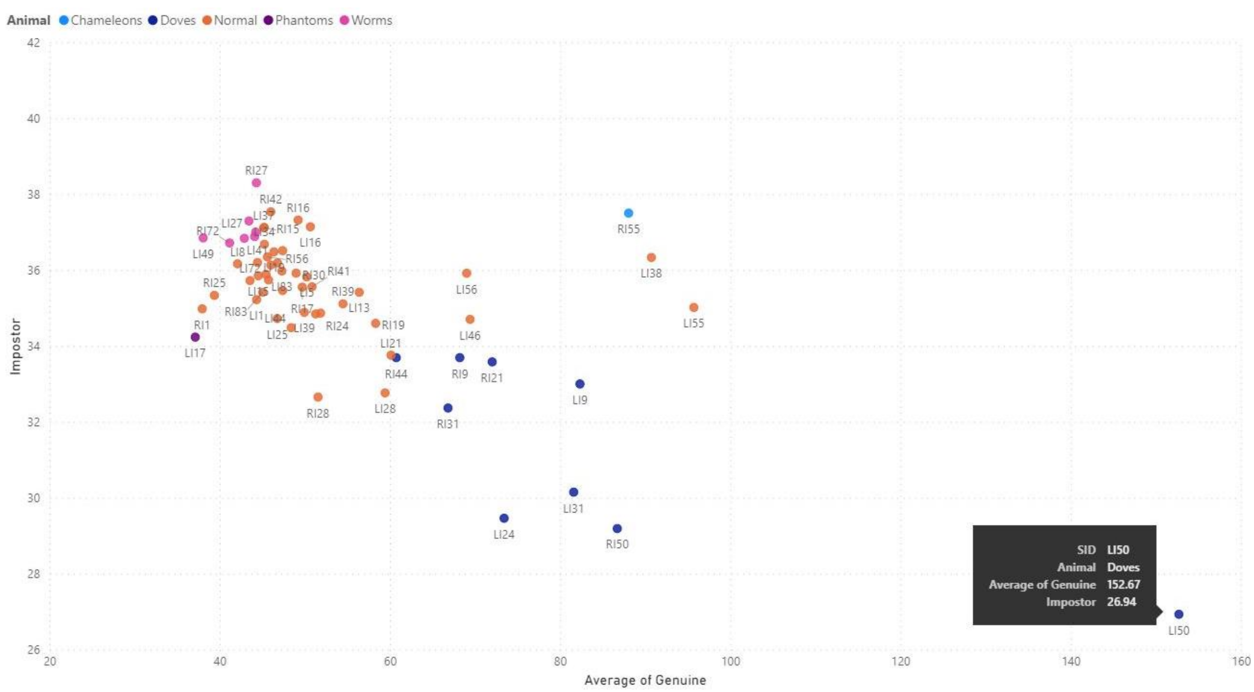

Figure 4.15 Zoo Plot: Subject LI50 at 7-12 Months Old

In *Figure 4.14*, while in the 0-6 months age group, the left index finger of subject 50 had an average genuine match score of 64.78 and an average impostor match score of 35.24. In *Figure 4.15*, while in the 7-12 months age group, the left index finger of subject 50 had an average genuine match score of 152.67 and an average impostor match score of 26.94. The larger change in match scores across age groups resulted in an SSI of 0.204. In comparison, the mean and median SSI for the 0-6 months and 7-12 months overlap was 0.031 and 0.018, respectively. This represented instability in subject 50's individual fingerprint performance.

4.4.2.2 Overlap Between 7-12 Months and >12 Months

Figures 4.16-4.21 display the zoo analysis for individuals who were in both the 7-12 months and the >12 months age groups *Figure 4.16* shows the zoo analysis on overlapping subjects in the 7-12 months old group and *Figure 4.17* shows the zoo analysis on overlapping subjects in the >12 months old group. *Table 4.7* displays the overall animal classification breakdown by age group for subjects who were in both the 7-12 months and >12 months groups.

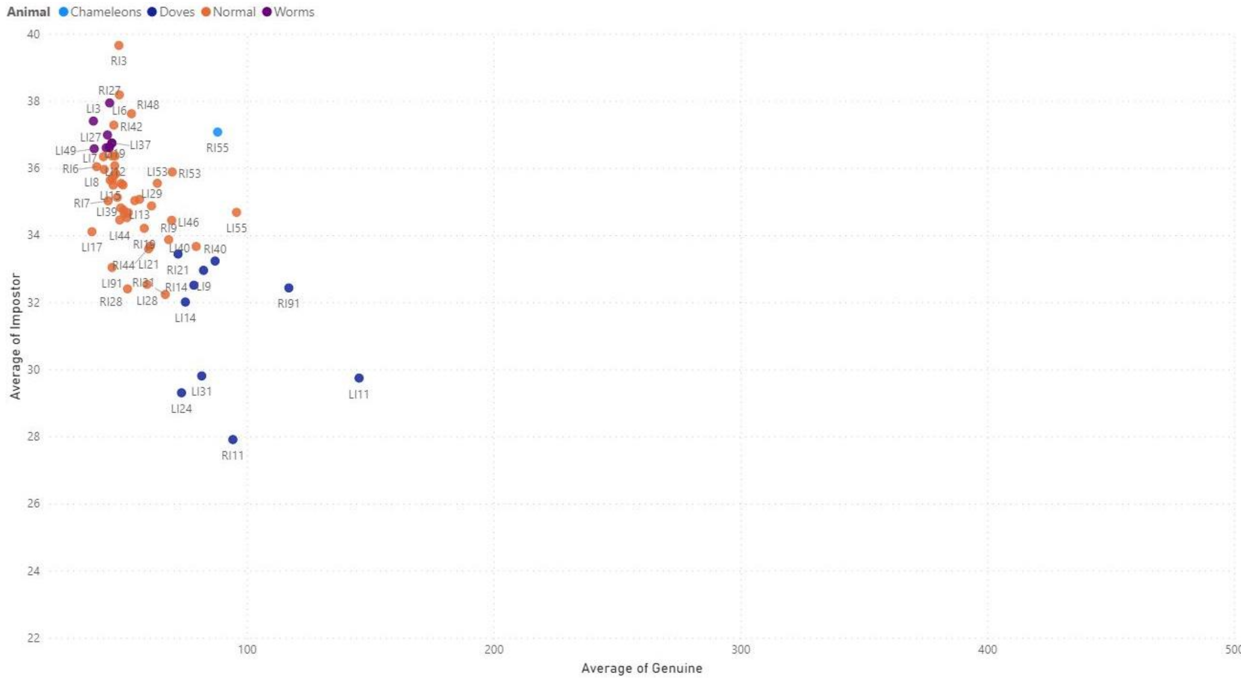

Figure 4.16 Zoo Plot: 7-12 Months Old (7-12 Months and >12 Months Overlap)

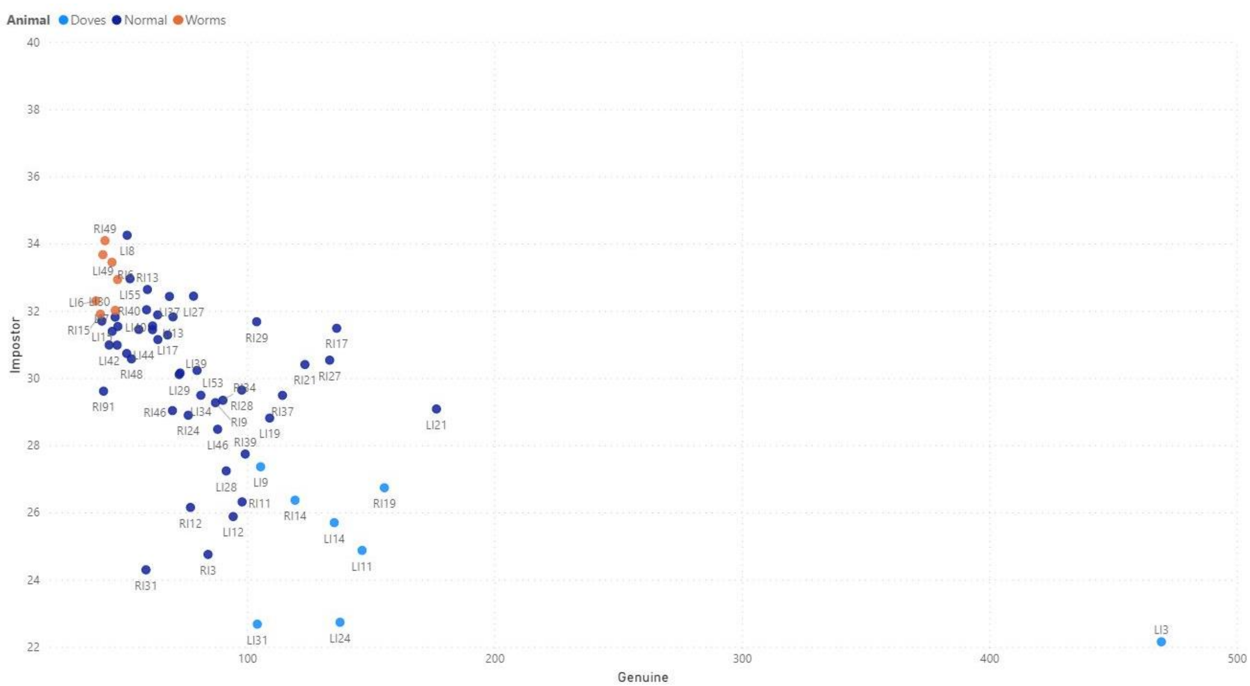

Figure 4.17 Zoo Plot: >12 Months Old (7-12 Months and >12 Months Overlap)

Table 4.7 Animal Classification Breakdown: 7-12 and >12 Months Overlap

Animal Classification	7-12 months	>12 months
Chameleon	1	0
Dove	10	8
Normal	42	47
Worm	9	7
Total	62	62

Overall, subjects were unstable across the 7-12 months and >12 months age groups. It was interesting to note that there was an increase in normal classifications and a decrease in both dove and worm classifications, as subjects from the 7-12 months group aged into the >12 months group. This means that while the DET curves showed improvement in EER between 7-12 months and >12 months, some of these individuals performed worse with respect to the >12 months population than they did in comparison to the 7-12 months old group.

4.4.2.2.1   Stability Between 7-12 Months and >12 Months

Figures 4.18 and *4.19* illustrate an example of individual stability across the 7-12 months and >12 months age groups.

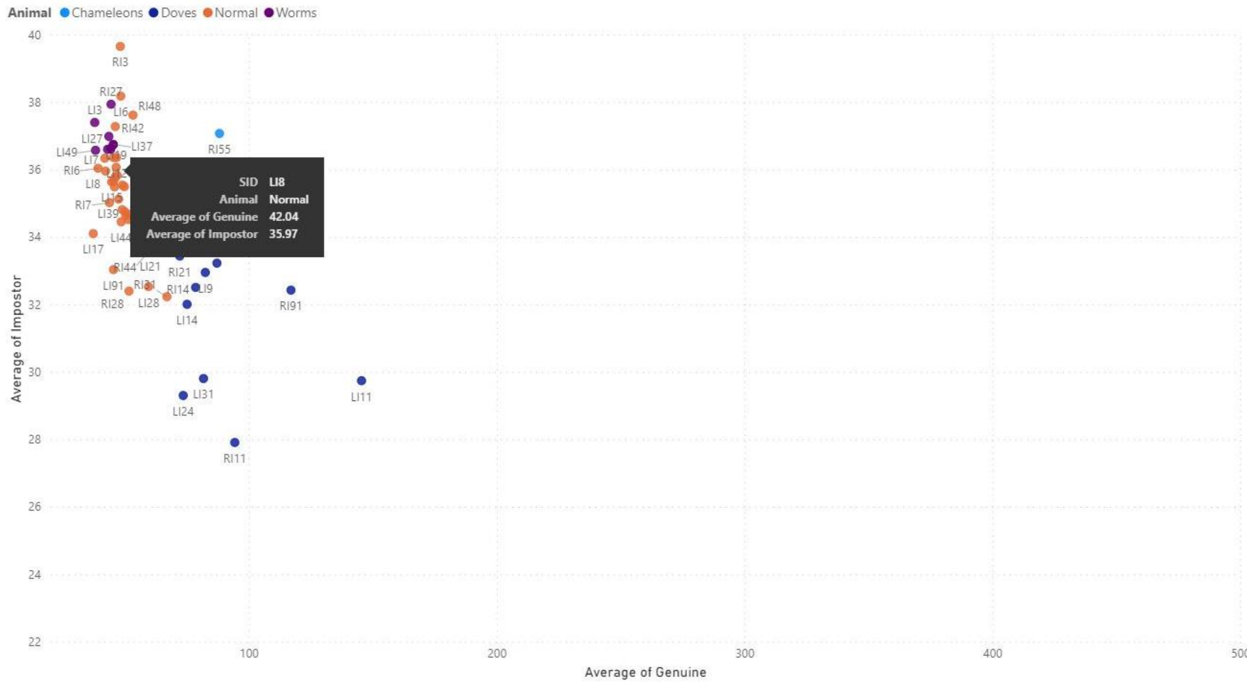

Figure 4.18 Zoo Plot: Subject LI8 at 7-12 Months Old

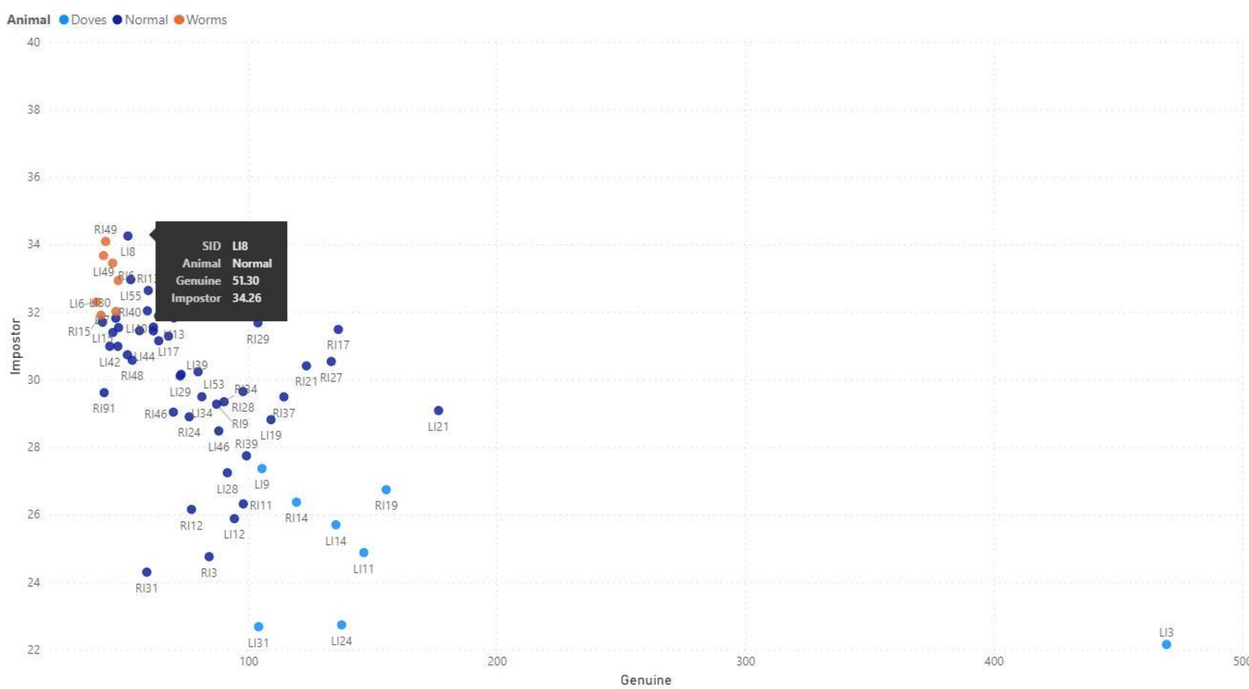

Figure 4.19 Zoo Plot: Subject LI8 at >12 Months Old

In *Figure 4.18*, while in the 7-12 months age group, the left index finger of subject 8 had an average genuine match score of 42.04 and an average impostor match score of 35.97. In *Figure 4.19*, while in the >12 months age group, the left index finger of subject 8 had an average genuine match score of 51.30 and an average impostor match score of 34.26. This minimal change in match scores across age groups resulted in an SSI of 0.022. In comparison, the mean and median SSI for the 7-12 months and >12 months overlap was 0.081 and 0.049, respectively. This represented stability in subject 8's individual fingerprint performance.

4.4.2.2.2 Instability Between 7-12 Months and >12 Months

Figures 4.20 and *4.21* illustrate an example of individual instability across the 7-12 months and >12 months age groups

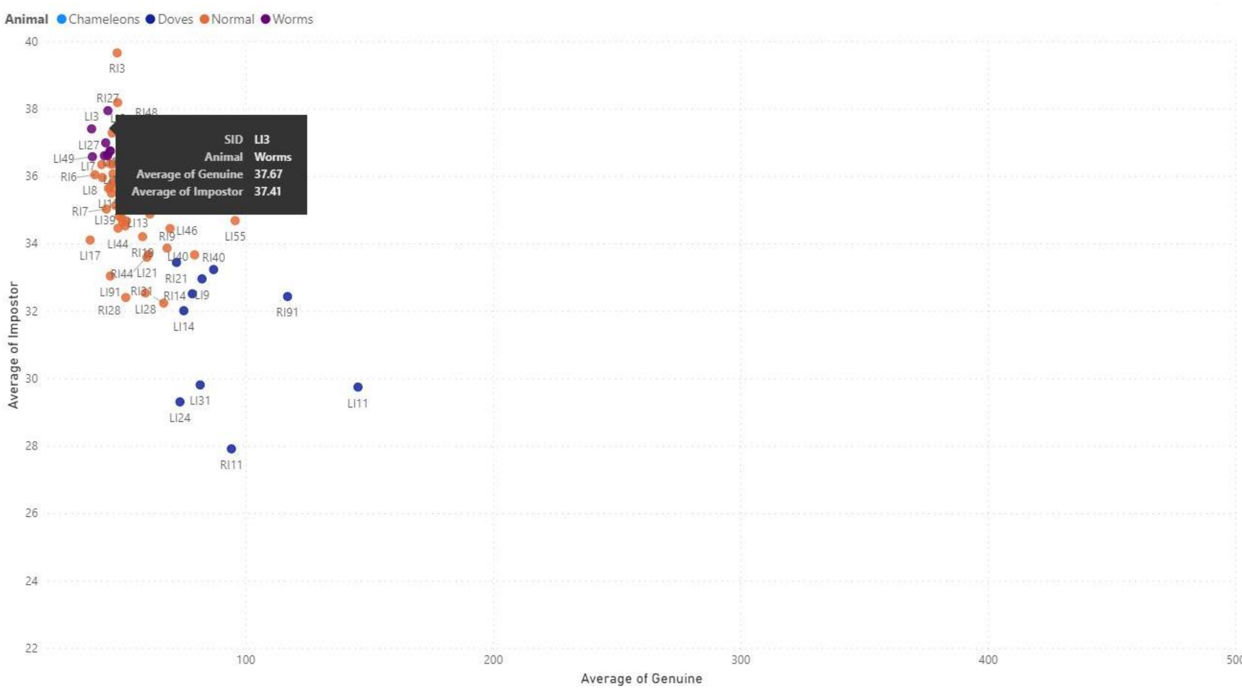

Figure 4.20 Zoo Plot: Subject LI3 at 7-12 Months Old

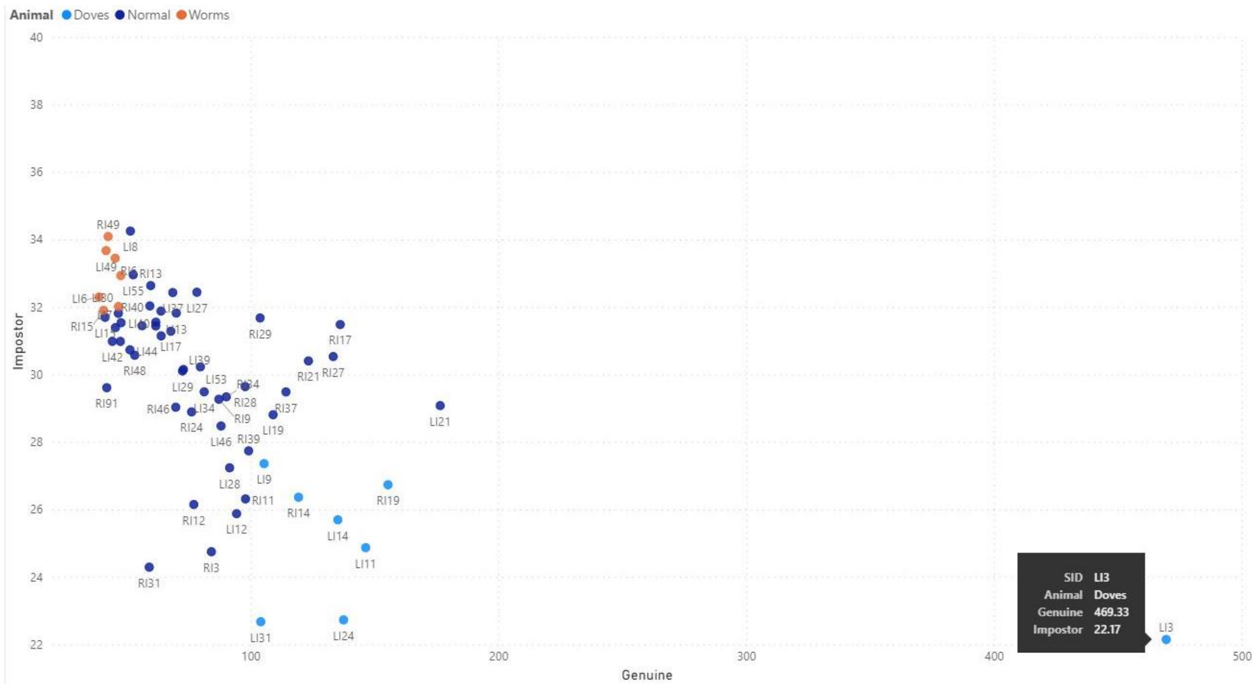

Figure 4.21 Zoo Plot: Subject LI3 at >12 Months Old

In *Figure 4.20*, while in the 7-12 months age group, the left index finger of subject 3 had an average genuine match score of 37.67 and an average impostor match score of 37.41. In *Figure 4.21*, while in the >12 months age group, the left index finger of subject 3 had an average genuine match score of 469.33 and an average impostor match score of 22.17. The larger change in match scores across age groups resulted in an SSI of 0.998. In comparison, the mean and median SSI for the 7-12 months and >12 months overlap was 0.081 and 0.049, respectively. This represented instability in subject 3’s individual fingerprint performance.

4.4.2.3   Overlap Between the 0-6, 7-12, and >12 Months Age Groups

Figures 4.22-4.30 display the zoo analysis for individuals who were in all three age groups. *Figure 4.22* shows the zoo analysis on overlapping subjects in the 0-6 months old group, *Figure 4.23* shows the zoo analysis on overlapping subjects in the 7-12 months old group, and *Figure 4.24* shows the zoo analysis on overlapping subjects in the >12 months old group. *Table*

4.8 displays the overall animal classification breakdown by age group for subjects who were in all three age groups.

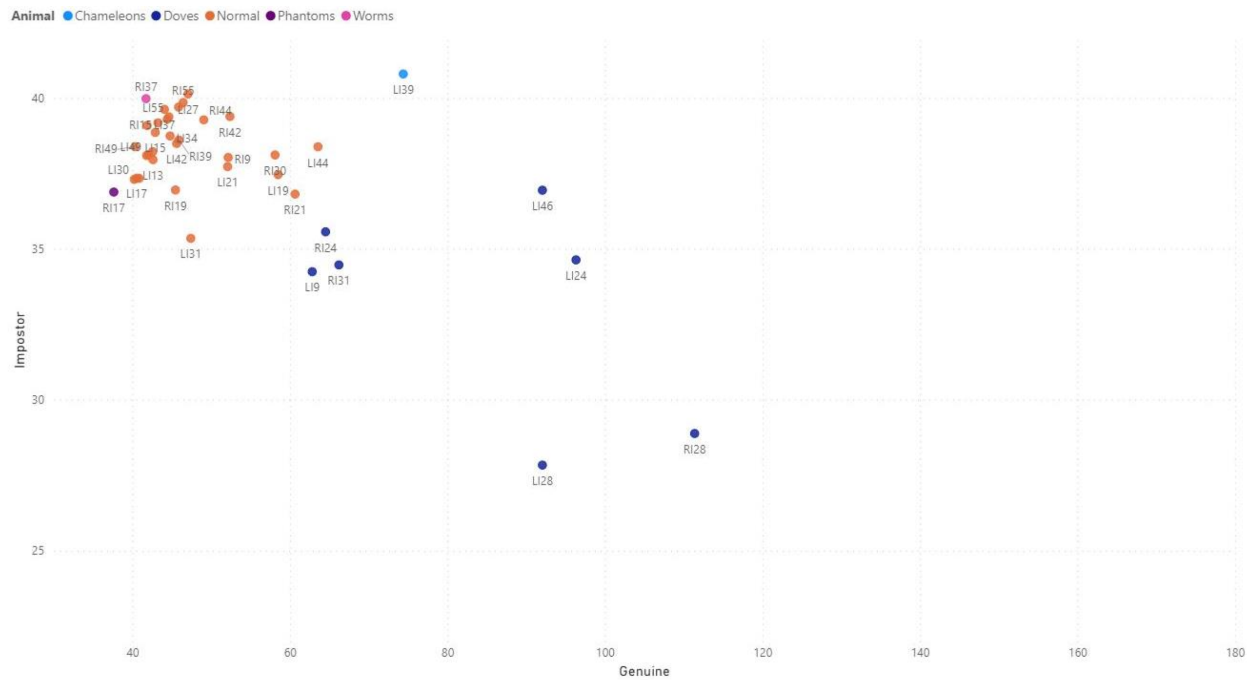

Figure 4.22 Zoo Plot: 0-6 Months Old (Overlap Between All Three Groups)

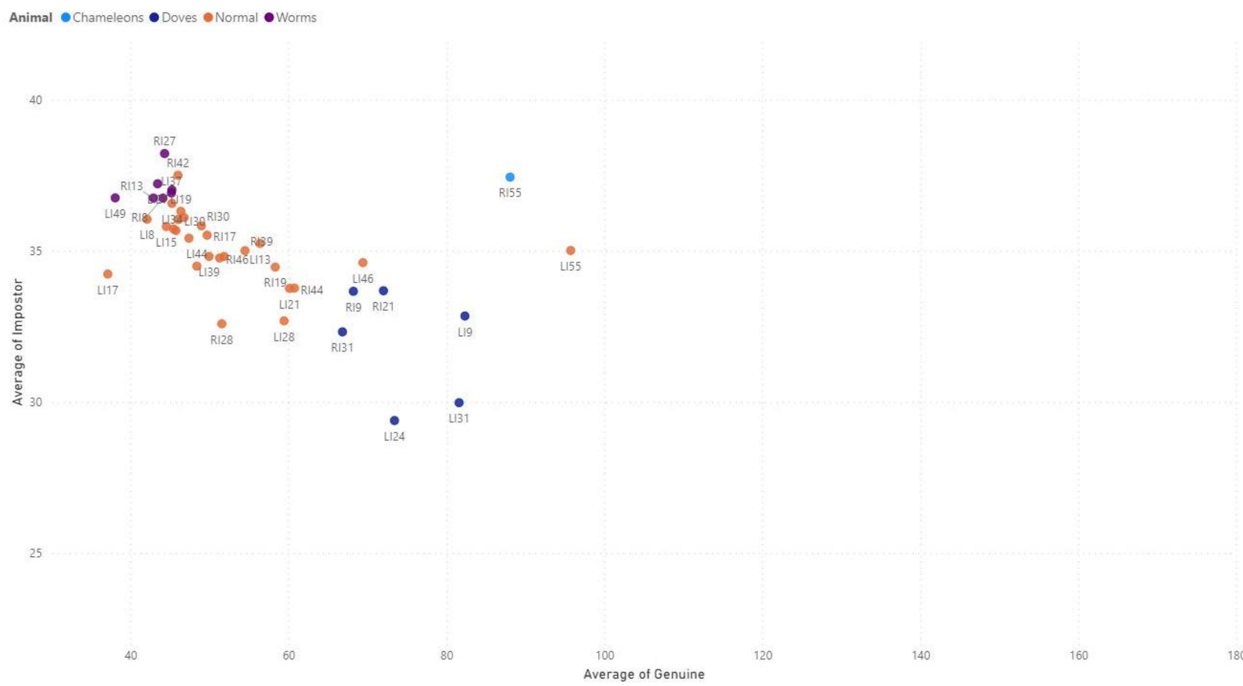

Figure 4.23 Zoo Plot: 7-12 Months Old (Overlap Between All Three Groups)

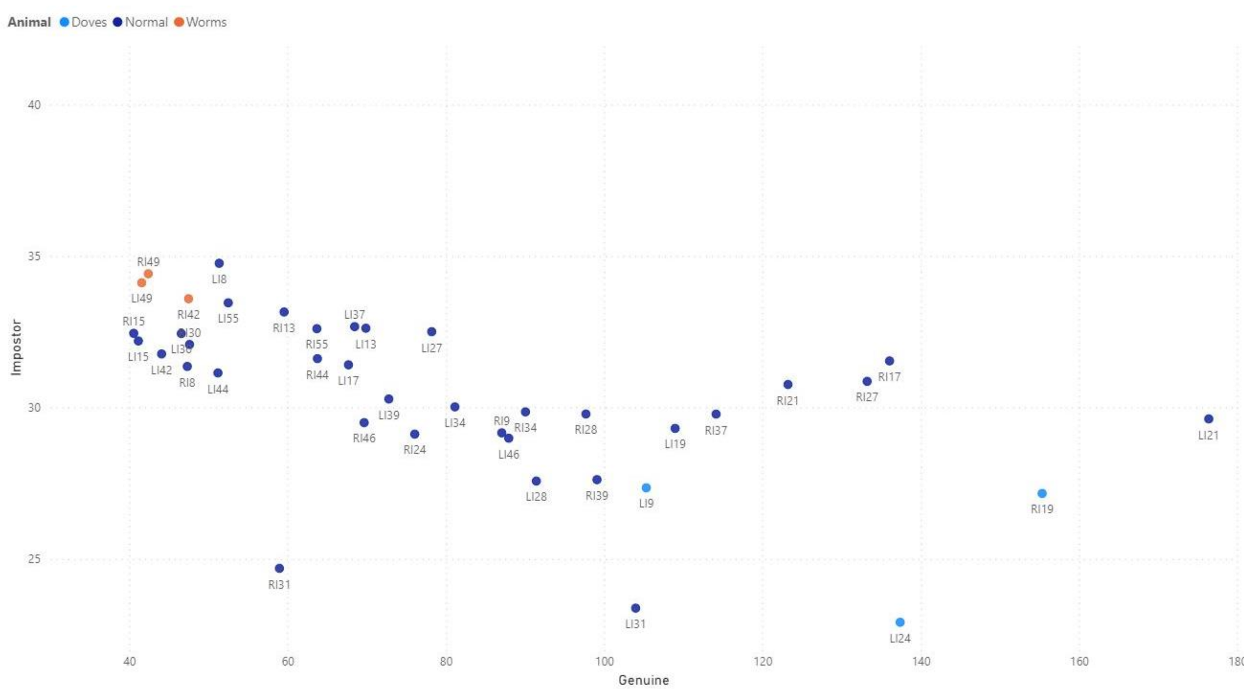

Figure 4.24 Zoo Plot: >12 Months Old (Overlap Between All Three Groups)

Table 4.8 Animal Classification Breakdown: Overlap Between All Three Groups

Animal Classification	0-6 months	7-12 months	>12 months
Chameleon	1	1	0
Dove	7	6	3
Normal	30	26	34
Phantom	1	0	0
Worm	1	7	3
Total	40	40	40

Overall, subjects were unstable across all three age groups. There was less stability across all groups than in either of the other two previously evaluated group overlaps. There were also some instances of animal classifications flipping back and forth. As age increased, the number of dove classifications decreased. It was interesting to note that there was a decrease in normal classifications between 0-6 months and 7-12 months, and then an increase in normal classifications between 7-12 months and >12 months. There was also an increase in worm classifications between 0-6 months and 7-12 months, and then a decrease in worm classifications between 7-12 months and >12 months.

4.4.2.3.1 Stability Between All Three Groups

Figures 4.25-4.27 illustrate an example of individual stability across all three age groups.

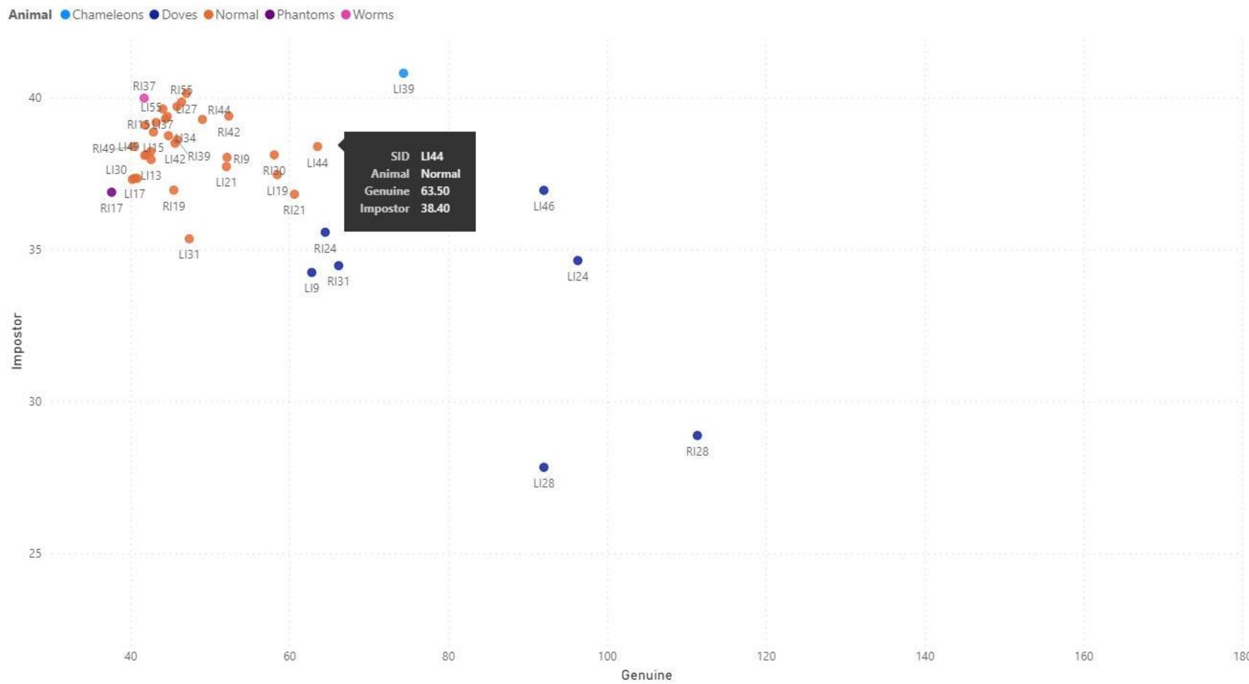

Figure 4.25 Zoo Plot: Subject LI44 at 0-6 Months Old

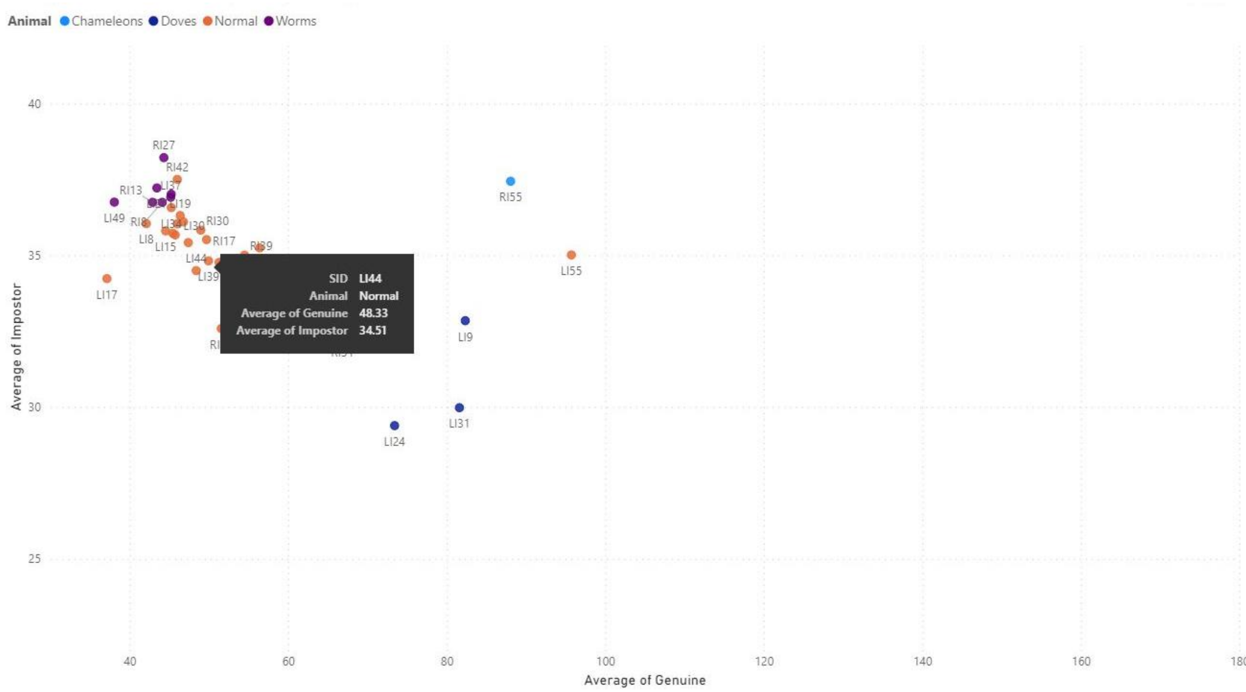

Figure 4.26 Zoo Plot: Subject LI44 at 7-12 Months Old

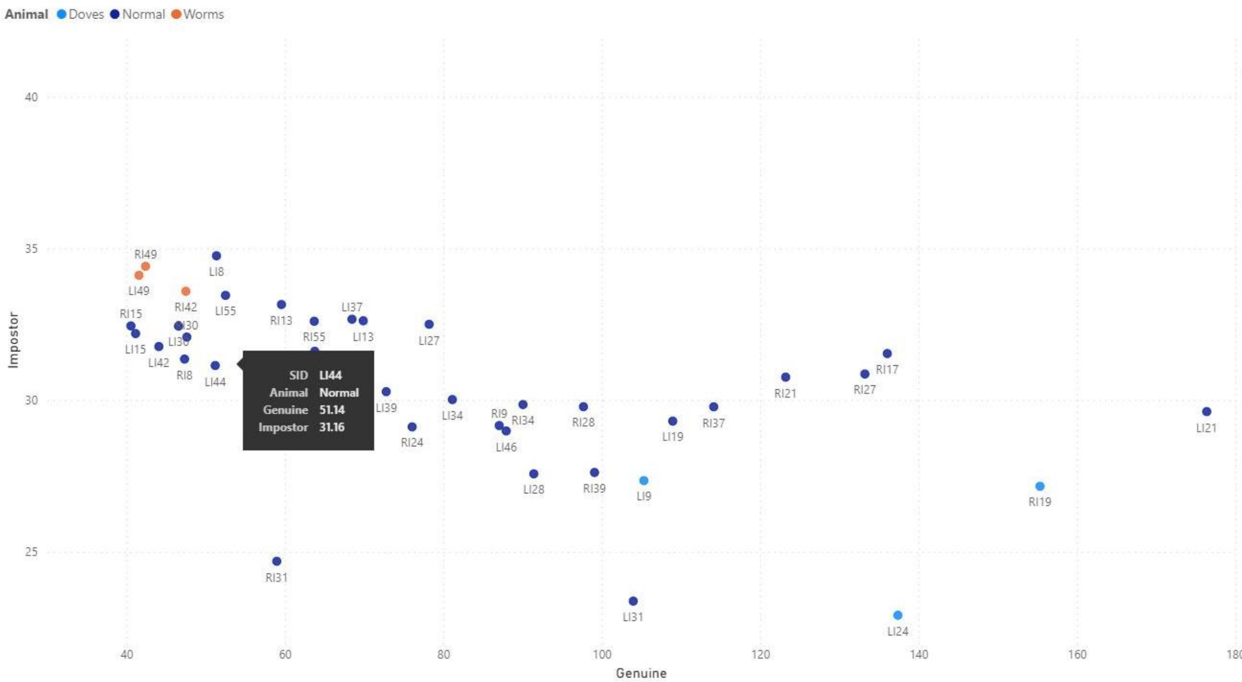

Figure 4.27 Zoo Plot: Subject LI44 at >12 Months Old

In *Figure 4.25*, while in the 0-6 months age group, the left index finger of subject 44 had an average genuine match score of 63.50 and an average impostor match score of 38.40. In *Figure 4.26*, while in the 7-12 months age group, the left index finger of subject 44 had an average genuine match score of 48.33 and an average impostor match score of 34.51. In *Figure 4.27*, while in the >12 months age group, the left index finger of subject 44 had an average genuine match score of 51.14 and an average impostor match score of 31.16. This minimal change in match scores across age groups represented stability in subject 44's individual fingerprint performance. The SSI values for subject LI44 between all age groups are displayed below in *Table 4.9*.

Table 4.9 SSI: Subject LI44

Age Group	G1-G2	G2-G3	G1-G3
LI44	0.036	0.010	0.033
Overall Mean	0.031	0.073	0.074

4.4.2.3.2 Instability Between All Three Age Groups

Figures 4.28-4.30 illustrate an example of individual instability across all age groups.

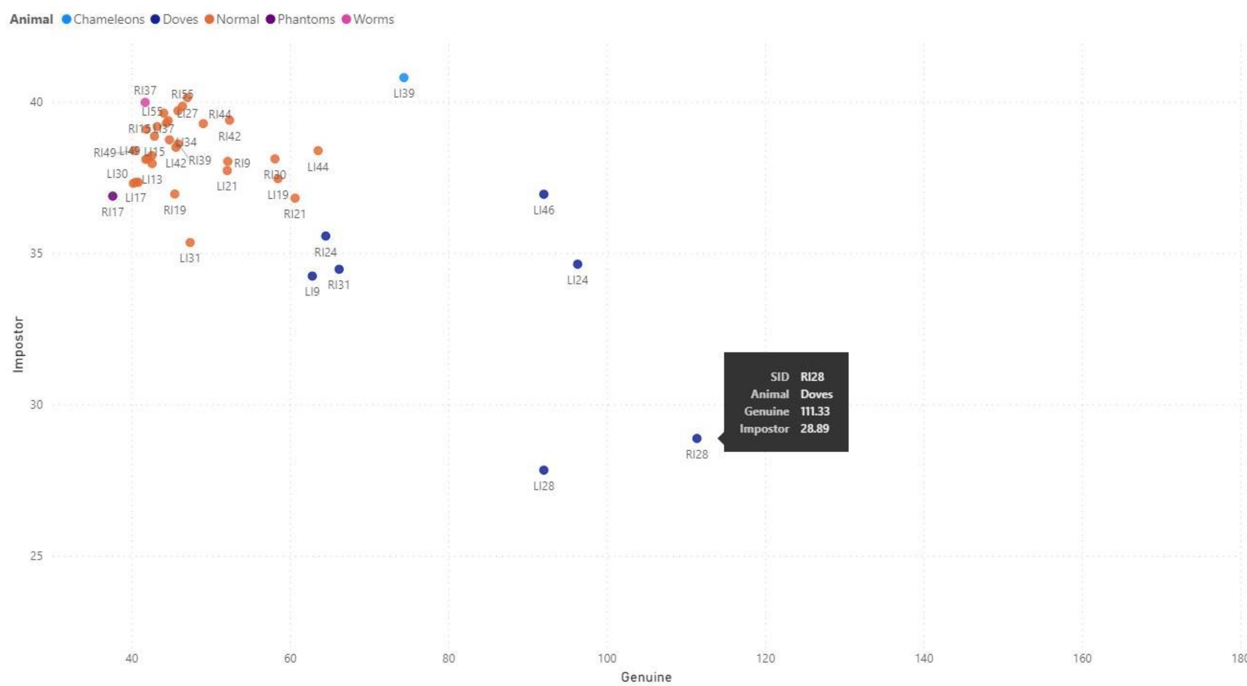

Figure 4.28 Zoo Plot: Subject RI28 at 0-6 Months Old

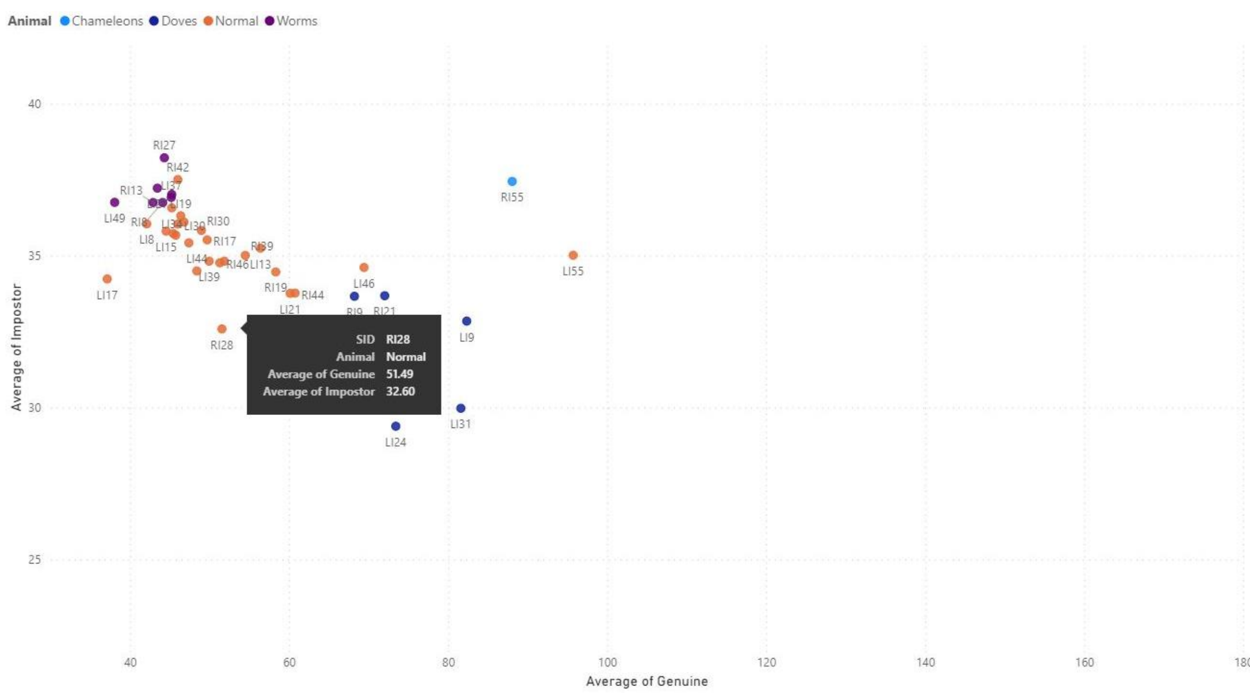

Figure 4.29 Zoo Plot: Subject RI28 at 7-12 Months Old

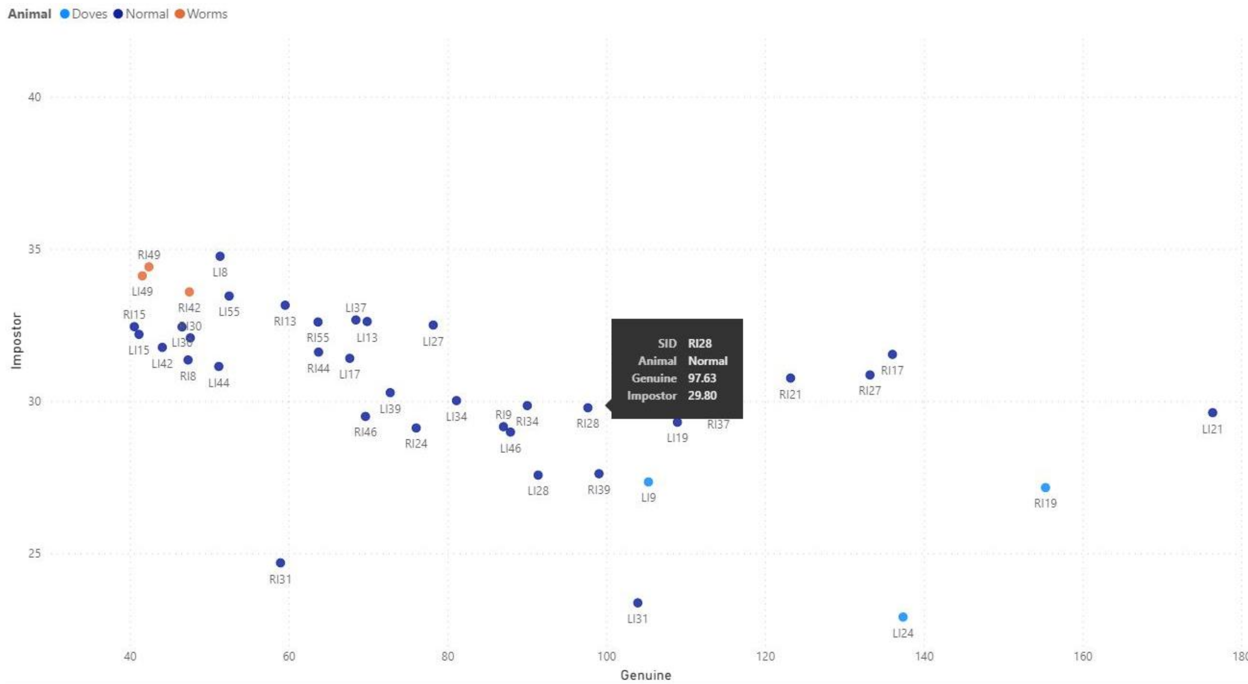

Figure 4.30 Zoo Plot: Subject RI28 at >12 Months Old

In *Figure 4.28*, while in the 0-6 months age group, the right index finger of subject 28 had an average genuine match score of 111.33 and an average impostor match score of 28.89. In *Figure 4.29*, while in the 7-12 months age group, the right index finger of subject 28 had an average genuine match score of 51.49 and an average impostor match score of 32.60. In *Figure 4.30*, while in the >12 months age group, the right index finger of subject 28 had an average genuine match score of 97.63 and an average impostor match score of 29.80. This large change in match scores across age groups represented stability in subject 28's individual fingerprint performance. The SSI values for subject RI28 between all age groups are displayed below in *Table 4.10*.

Table 4.10 SSI: Subject RI28

Age Group	G1-G2	G2-G3	G1-G3
RI28	0.139	0.107	0.032
Overall Mean	0.031	0.073	0.074

4.4.2.4 Overall Stability Score Index

The stability score index (SSI) was used to evaluate the stability of each individual as they aged into different age groups during the data collection process. *Tables 4.11*, displayed below, illustrates the overall mean and median SSI between age groups for children who overlapped between each of those two groups. *Table 4.12* shows the SSI by age group for children who overlapped between all three age groups. *Appendix C* illustrates the SSI for each child. Group 1 refers to the 0-6 months group, group 2 refers to the 7-12 months group, and group 3 refers to the >12 months old group.

Table 4.11 SSI: Overlap Between Each Age Group

	G1-G2	G2-G3
Mean	0.0308	0.0814
Median	0.0180	0.0494

Table 4.12 SSI: Overlap Between All Three Groups

	G1-G2	G2-G3	G1-G3
Mean	0.0313	0.0726	0.0742
Median	0.0216	0.0543	0.0427

In *Table 4.11*, both the mean and median SSI values increase from the 0-6 months & 7-12 months comparison to the 7-12 months & >12 months comparison. This represented an increase in instability, or a larger change in individual performance. In *Table 4.12*, both the mean and median SSI values exhibited the same increase from 0-6 & 7-12 months to 7-12 & >12 months. This represented the same increase in instability. However, the SSI values did not change between the 7-12 months & >12 months and 0-6 months & >12 months comparison. This represented stability between both comparisons, and further illustrated that there was no difference in performance between the 0-6 months group and 7-12 months group. The difference between the 7-12 months group and >12 months group was most responsible for the difference in individual performance, following the same trend found earlier when evaluating overall system performance. Overall, there was individual instability across age groups which likely impacted system performance and influenced the DET curves.

CHAPTER 5. CONCLUSIONS AND FUTURE WORK

This chapter is divided into two sections. The first section discusses the conclusions regarding the hypotheses of this study. The second section discusses future work to be done in infant fingerprint recognition, and recommendations based on the findings of this study and items that this study did not cover.

5.1 Conclusions

One conclusion that was drawn from this research was image quality and minutiae count are impacted by age. There was a significant difference in quality metrics based upon a child's age, F (4, 5412) = 306.20, $p < .001$; Wilk's $\Lambda = .67$, partial $\eta^2 = .19$. Age group had a statistically significant effect on both image quality (F (2, 2707) = 661.47; $p < .001$; partial $\eta^2 = .33$) and minutiae count (F (2, 2707) = 152.88; $p < .001$; partial $\eta^2 = .10$). The Games-Howell test showed that image quality was significantly different between all age group comparisons ($p < .001$) and the Tukey test showed that minutiae count was significantly different between all age group comparisons ($p < .001$).

Additionally, analysis of DET curves, zoo characteristics, and the stability score index all showed a larger change between the >12 months group and all other age groups, meaning that 12 months old serves as an intriguing age to focus on for further research and the development and testing of new technology. In accordance with Jain et al. (2017), 12 months old appeared to be a benchmark age for using fingerprint recognition on children. However, there was a much larger FRR at the same FAR found in this study than was found in Galbally (2019). This could be a result of this study including more subjects who were less than 12 months old and the minutiae count threshold being turned off for this study's data collection. *Table 4.5* compares the DET curves (FAR and FRR) of this study, Reiff (2020), with the performance of children (0-4 years old), adults (18-25 years old), and the elderly (65-98 years old) that was found in Galbally (2019).

Table 5.1 FAR and FRR: Comparison with Previous Study by Age Group

Study	Age (Years)	FAR (%)	FRR (%)
Galbally (2019)	0-4	0.1	37.00
	18-25	0.1	1.50
	65-98	0.1	6.50
Reiff (2020)	0-5	0.1	73.15

Another conclusion that can be drawn is that infant fingerprint recognition performance increased with age. As the age group got older the EER rates decreased from 36.00% to 14.25%, indicating less matching error within the system as age was increased. FTX rates decreased from 8.89% to 0% as age increased. By 7 months old, the FTX rates decreased to a level 0.86%. The genuine match scores increased from 50.76 to 123.96 as age increased, and impostor match scores decreased from 37.97 to 25.53 as age increased. This means that matching performance improved with age. Again, this supports 12 months old as an intriguing age for the research of infant fingerprint recognition.

While there was a difference in performance between age groups, there was generally stability for subjects who overlapped between multiple age groups. It was concluded based upon this study, that individual instability did not significantly impact overall performance. Difference in performance would most likely be attributed to the difference in physical characteristics between subjects in each age group.

According to the Dutch Ministry of the Interior and Kingdom Relations (2005), infant fingerprints appear unviable for fingerprint recognition purposes, However, from this study it was concluded that 12 months old appears to be a turning point where fingerprint quality and matching performance improve. This serves as a potential benchmark for using fingerprint recognition with children.

5.2 Future Work

The dataset for this study was chosen for secondary analysis because Hutchison (2018) used the same multimodal infant biometric data collection for their secondary analysis on infant iris recognition. There it was recommended that "a comparison of the same subjects across different biometric modalities will help the biometric research community understand the most suitable biometrics for infants" (Hutchison, 2018, p. 70). In accordance with Hutchison (2018), one recommendation would be to conduct research that combines biometric modalities on the same children. This could help determine whether certain age groups perform better or worse, in general, as well as if poor performance in certain individuals is due to their specific behavioral or physical traits. This could also be used to analyze if a capture method or additional modality could be used to mitigate any of these challenges associated with infant fingerprint recognition and improve the overall biometric system performance. Examples of this would in adult biometric recognition include the use of face and iris multimodal recognition systems, where an image of both is taken. Using both can help mitigate shortcomings of either modality by itself. It would be interesting to see what could work in this manner for infant biometrics.

Another recommendation would be to collect more infant fingerprints and over a longer period. There has not been a lot of promising test results for infant fingerprint recognition, using the technologies we currently have available to us. Galbally (2019) discovered that fingerprint quality and performance increased with age for children, peaked at the young adult stage, and then proceeded to regress by the elderly stage. A lot of research has been conducted to analyze and interpret the decrease in performance as young adults age into the elderly population. However, it would also be important to analyze the increase in performance with age from children to adults. This could provide a more accurate view of the larger picture, regarding aging, and could lead to better utilization or development of biometric systems. Additionally, this study used the same subjects as Hutchison (2018), therefore a multimodal analysis to determine potential correlation between fingerprint and iris recognition on the same infants could also be conducted.

Infant interaction and behavior were not recorded and analyzed for this study. The understanding of these phenomena could help understand infant biometric recognition and where technology/system improvements can and cannot be made. It would be interesting to determine whether FTXs occurred as a product of the user or environment. It would be recommended to analyze whether children who produce an FTX consistently perform worse than others or if it was a random occurrence. Additionally, it would be interesting to know whether random or consistent FTXs exhibit any pattern or consistency with which attempt or which order that the FTX occurred. Pressure/force level would be another behavioral factor worth reviewing. Uneven pressure being applied when capturing infant fingerprint images may be contributing to quality issues. Ideal force level may be different for infant fingers than adult fingers due to factors such as size or chemical composition. Since parents/guardians may need to place the infant’s finger on the sensor, knowing the correct amount of applied force needed could help improve the quality of data being collected. Hand dominance is also another behavioral trait that could be further examined as children age and develop hand dominance.

Another recommendation is to test different sensors or sensor technologies. Certain behavioral or physical characteristics can be mitigated with sensors that use a different capture technology (optical, capacitive, etc.) or that have a different shape and size that may benefit the smaller fingers of infants.

Testing all ten fingers from each infant could allow us to determine whether other fingers are likely to be negatively affected by an infant's behavioral characteristics, or if better performance/capture methods can be achieved. Thumbs have a larger surface area than other fingers and more minutiae points may help mitigate some quality issues with infant fingerprints.

Moving forward, it is important to identify what qualifies as good quality, good performance, and the best practices to achieve this. As future research is conducted on infant fingerprint recognition and other infant biometrics, some of these guidelines can be considered and put into place. It is very difficult for researchers to make concrete conclusions until it is known what constitutes as success in a practical application or environment. This study adds to the body of knowledge and serves as a continuation of prior research. Future work will increase this knowledge and continue to add insight that the brackets around good quality and performance can be tightened and more clearly defined, for all modalities of infant biometrics. With this in place, it may be possible to design new technology or better use current systems for the purpose of infant biometric recognition.

REFERENCES

Adler, A., & Schuckers, M. E. (2007). Comparing Human and Automatic Face Recognition Performance. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, *37*(5), 1248–1255. doi: 10.1109/tsmcb.2007.907036

Akins, K. (n.d.). The case of the disappearing fingerprints. Retrieved December 7, 2019, from http://karlaakins.com/case-of-the-disappearing-fingerprints/.

Alonso-Fernandez, F., Veldhuis, R. N., Bazen, A. M., Fierrez-Aguilar, J., & Ortega-Garcia, J. (2006). Sensor Interoperability and Fusion in Fingerprint Verification: A Case Study Using Minutiae- and Ridge-based Matchers. Retrieved December 11, 2018.

AlShehri, H., Hussain, M., AboAlSamh, H., AlZuair, M. (2018). A Large-Scale Study of Fingerprint Matching Systems for Sensor Interoperability Problem. *Sensors*,*18*(4), 1008. doi:10.3390/s18041008

Bachenheimer, D. (2016). *UNHCR Biometric Identity Management System*. Conference presented at the International Biometric Performance Conference, Gaithersburg, MD.

Bharadwaj, S., Bhatt, H. S., Singh, R., Vatsa, M., & Singh, S. K. (2010). Face recognition for newborns: A preliminary study (pp. 1–6). IEEE. https://doi.org/10.1109/BTAS.2010.5634500

Bishop, M. (2003). *Computer Security*. Addison-Wesley.

Bowyer K.W., Hollingsworth K., Flynn P. J. (2007) ‘Image understanding for iris biometrics: A survey’. Computer Vision and Image Understanding, 110, pp 281– 307

Cadd, S., Islam, M., Manson, P., & Bleay, S. (2015). Fingerprint composition and aging: A literature review. *Science & Justice*,*55*(4), 219-238. doi:10.1016/j.scijus.2015.02.004

Carls, J. W. (2009). A Framework for Analyzing Biometric Template Aging and Renewal Prediction. *Air Force Institute of Technology*. Retrieved January 16, 2019.

Carls, J. W., Raines, R., Grimaila, M., & Rogers, S. (2008). Biometric enhancements: Template aging error score analysis. *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition*. doi: 10.1109/afgr.2008.4813331

Corby, P. M., Schleyer, T., Spallek, H., Hart, T. C., Weyant, R. J., Corby, A. L., & Bretz, W. A. (2006). Using Biometrics for Participant Identification in a Research Study: A Case Report. *Journal of the American Medical Informatics Association*, *13*(2), 233–235. https://doi.org/10.1197/jamia.M1793

Cummins, H. & Midlo, C. (1961). Finger prints, Palms and Soles: An Introduction to Dermatoglyphics. Dover Publications New York, 1961, vol. 319.

Doddington, G., Liggett, W., Martin, A., Przybocki, M., & Reynolds, D. (1998). Sheep, goats, lambs and wolves: A statistical analysis of speaker performance in the NIST 1998 speaker recognition evaluation. *The 5th International Conference on Spoken Language Processing, 13*, 1–5 (1998).

Dunstone, T., & Yager, N. (2009). *Biometric system and data analysis*. New York, NY: Springer Science+Business Media.

Dutch Ministry of the Interior and Kingdom Relations. (2005). *Evaluation Report Biometrics Trial 2b or not 2b* (Technical Report). Dutch Ministry of the Interior and Kingdom Relations.

Elliott, S., Kukula, E., & Sickler, N. (2007). The Challenges of the Environment and the Human / Biometric Device Interaction on Biometric System Performance

Fairhurst, M., Erbilek, M., & Costa-Abreu, M. D. (2015). Selective Review and Analysis of Aging Effects in Biometric System Implementation. *IEEE Transactions on Human-Machine Systems*, *45*(3), 294–303. doi: 10.1109/thms.2014.2376874

Fenker, S. P., & Bowyer, K. W. (2011). Experimental evidence of a template aging effect in iris biometrics. *2011 IEEE Workshop on Applications of Computer Vision (WACV)*. doi: 10.1109/wacv.2011.5711508

Galbally, J., Haraksim, R., & Beslay, L. (2019). A Study of Age and Ageing in Fingerprint Biometrics. *IEEE Transactions on Information Forensics and Security*, *14*(5), 1351–1365. doi: 10.1109/tifs.2018.2878160

Harvey, J., Campbell, J., Elliott, S., Brockly, M., & Adler, A. (2017). Biometric Permanence: Definition and Robust Calculation. *2017 Annual IEEE International Systems Conference (SysCon)*. doi: 10.1109/syscon.2017.7934760

Hong, L., Wan, Y., & Jain, A. (1998). Fingerprint image enhancement: Algorithm and performance evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence,20*(8), 777-789. doi:10.1109/34.709565

Hutchison, T. (2018). Longitudinal Analysis on the Feasibility of Iris Recognition Performance For Infants 0-2 Years Old. Purdue University

ISO/IEC JTC 1 SC 37. (2005). Text of FCD 19795-1, Biometric Performance Testing and Reporting - Part 1: Principles and Framework. (N908). ISO/IEC. Geneva.

ISO/IEC 2382-37. (2017). Information Technology – Vocabulary – Part 37: Biometrics. ISO/IEC. Geneva.

ISO/IEC TR 24741. (2006). Information Technology - Biometrics Tutorial. ISO/IEC. Geneva.

Jagadeesan, A., Thillaikkarasi, T., & Duraiswamy, D. (2010). Cryptographic Key Generation from Multiple Biometric Modalities: Fusing Minutiae with Iris Feature. *International Journal of Computer Applications*,*2*(6), 16-26. doi:10.5120/673-946

Jain, A. (2005). Biometric Recognition: How Do I Know Who You Are? *Springer*

Jain, A., Arora, S., Best-Rowden, L., Cao, K., Sudhish, P. S., & Bhatnagar, A. (2015). *Biometrics for Child Vaccination and Welfare: Persistence of Fingerprint Recognition for Infants and Toddlers* (Technical Report No. MSU-CSE-15-7) (p. 17). Michigan State University.

Jain, A., Arora, S., Cao, K., Best-Rowden, L., & Bhatnagar, A. (2016). Fingerprint Recognition of Young Children. *IEEE Technical Report*.

Jain, A., Arora, S., Cao, K., Best-Rowden, L., & Bhatnagar, A. (2017). Fingerprint Recognition of Young Children. *IEEE Transactions on Information Forensics and Security*, *12*(7), 1501–1514. https://doi.org/10.1109/TIFS.2016.2639346

Jain, A., Bolle, R., & Pankanti, S. (2002). Introduction to biometrics. (Jain, A., Bolle, R., and Pankanti, S.) *Biometrics: Personal Identification in Networked Society* (pp. 1-41). Springer. doi:10.1007/0-306-47044-6_1

Jain, A., Cao, K., & Arora, S. S. (2014). Recognizing infants and toddlers using fingerprints: Increasing the vaccination coverage (pp. 1–8). IEEE. https://doi.org/10.1109/BTAS.2014.6996252

Jain, A., & Kumar, A. (2010). Biometrics of Next Generation: An Overview. *Springer: Second Generation Biometrics.*

Jain, A., Ross, A., & Prabhakar, S. (2004). An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology,14*(1), (pp. 4-20). doi:10.1109/tcsvt.2003.818349

Jia, W., Cai, H.-Y., Gui, J., Hu, R.-X., Lei, Y.-K., & Wang, X.-F. (2012). Newborn footprint recognition using orientation feature. *Neural Computing and Applications*, *21*(8), 1855–1863. https://doi.org/10.1007/s00521-011-0530-9 82

Jia, W., Hu, R.-X., Gui, J., & Lei, Y.-K. (2010). Newborn Footprint Recognition Using Band-Limited Phase-Only Correlation. In D. Zhang & M. Sonka (Eds.), *Medical Biometrics* (Vol. 6165, pp. 83–93). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-13923-9_9

Khan, A.I. (2015). Comparing and Improving Existing Fingerprint Recognition Algorithms. *Research Gate*.

Lanitis, A. (2010). A survey of the effects of aging on biometric identity verification. *International Journal of Biometrics*, *2*, 34-52. doi: 10.1504/ijbm.2010.030415

Ling, H., Soatto, S., Ramanathan, N., & Jacobs, D.W. (2007). A study of face recognition as people age. *Proc. IEEE 11th Int. Conf. Comput. Vis.*, 1-8

Lugini, L., Marasco, E., Cukic, B., & Gashi, I. (2013). Interoperability between fingerprint biometric systems: An empirical study. doi:10.1109/dsnw.2013.6615516

LumenWaymaker. (2012). Introduction to Psychology. Retrieved December 9, 2019, from https://courses.lumenlearning.com/waymaker-psychology/chapter/reading-childhood/.

Makinde, O.A. (2015). Infant Trafficking and Baby Factories: A New Tale of Child Abuse in Nigeria. *Child Abuse Review*. doi.org/10.1002/car.2420

Mansfield, A.J., & Wayman, J.L. (2012). Best Practices in Testing and Reporting Performance of Biometric Devices.

Mason, S., Gashi, I., Lugini, L., Marasco, E., & Cukic, B. (2014). Interoperability between Fingerprint Biometric Systems: An Empirical Study. *2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*. doi:10.1109/dsn.2014.60

Michalski, D., Yiu, S. Y., & Malec, C. (2018). The Impact of Age and Threshold Variation on Facial Recognition Algorithm Performance Using Images of Children. *2018 International Conference on Biometrics (ICB)* (pp. 217-224). doi: 10.1109/icb2018.2018.00041

Moore, Z. E., Elliott, S. J., & O'Connor, K. J. (2015). Interoperability and the Stability Score Index. *IT in Industry*, *3*(4).

Muramoto, S. (2015). Strategies for Potential Age Dating of Fingerprints through the Diffusion of Sebum Molecules on a Nonporous Surface Analyzed using Time-of-Flight Secondary Ion Mass Spectrometry. *Analytical Chemistry*.

Neurotechnology. (n.d.). DigitalPersona U.are.U 4500 Scanner Specifications. Retrieved May 26, 2020, from https://www.neurotechnology.com/fingerprint-scanner-digitalpersona-u-are-u-4500.html

Newman, D. (2016, November 29). Top 10 Trends for Digital Transformation In 2017. Retrieved January 20, 2019, from https://www.forbes.com/sites/danielnewman/2016/08/30/top-10-trends-for-digital-transformation-in-2017/

O'Connor, K. J. (2013). Examination of Stability in Fingerprint Recognition across Force Levels. *Purdue University: Purdue E-Pubs*. Retrieved January 16, 2019.

O'Connor, L. (2016). This Is What It Was Like To Go To The Airport Before 9/11. Retrieved December 8, 2019, from https://www.huffpost.com/entry/airports-before-911_n_57c85e17e4b078581f11a133.

Patel, S., Bhattacharya, A., & Kumar, P. (2013). A Novel Approach To Classification Based Fingerprint Verification System. *International Journal of Scientific & Engineering Research, Vol. 4*.

Petrelli, M. (2009). The Impact of Finger Force on Single and Ten Print Fingerprint Recognition. Purdue University

Petry, B. (2015). The stability of the iris as a biometric modality. Purdue University.

Rebera, A., & Mordini, E. (2013). Biometrics and ageing: Social and ethical considerations. *Age Factors in Biometric Processing*. doi:10.1049/pbsp010e_ch3

Ross, A., & Jain, A. (2004). Biometric Sensor Interoperability: A Case Study in Fingerprints. *Biometric Authentication Lecture Notes in Computer Science,*134-145. doi:10.1007/978-3-540-25976-3_13

Ryu, J., Jang, J., & Kim, H. (2007). Analysis of Effect of Fingerprint Sample Quality in Template Ageing. *NIST Biometric Quality Workshop*, *II*.

Senseon Secure Access. (2020). Types of Access Control Systems & Why it is Important in 2020. Retrieved June 28, 2020, from https://www.senseon.com/access-control/

Shuckers, M. (2010). Additional topics and discussion. In M. Shuckers (Eds.) *Computational Methods in Biometric Authentication* (pp. 299-300). New York: Springer.

Sickler, N. & Elliott, S. (2005). An Evaluation of Fingerprint Quality across an Elderly Population vis-à-vis 18-25 year olds.

Tabassi, E. (2010). Image specific error rate: A biometric performance metric. *2010 20th International Conference on Pattern Recognition*, 1124–1127, doi:10.1109/ICPR.2010.281

Teli, M.N., Beveridge, J.R., Phillips, P.J., Givens, G.H., Bolme, D.S., & Draper, B.A. (2011). Biometric zoos: Theory and experimental evidence. *2011 International Joint Conference on Biometrics (IJCB)*, 1-8. doi:10.1109/IJCB.2011.6117479

Tiwari, S., Singh, A., & Singh, S. (2011). Newborn's ear recognition: Can it be done? *2011 International Conference on Image Information Processing* (pp. 1–6). https://doi.org/10.1109/ICIIP.2011.6108944

Topol, T. (2018, March 19). A comprehensive Timeline of 150 Years History in Biometrics. Passport-collector.com. Retrieved December 9, 2019, from https://www.passport-collector.com/a-comprehensive-timeline-of-150-years-history-in-biometrics/.

UNICEF USA. (2019). Child Protection and Development. https://www.unicefusa.org/mission/protect/trafficking. Retrieved May 29, 2020

Villemez, J. (2011, September 14). 9/11 to now: Ways we have changed. Retrieved December 8, 2019, from https://www.pbs.org/newshour/world/911-to-now-ways-we-have-changed.

Violino, B. (2015, March 3). Biometric security is on the rise. https://www.csoonline.com/article/2891475/biometric-security-is-on-the-rise.html. Retrieved December 8, 2019

Wayman, J., Jain, A. K., Maltoni, D., & Maio, D. (2005). An Introduction to Biometric Authentication Systems. In J. Wayman, A. Jain, D. Maltoni, & D. Maio (Eds.), Biometric Systems (pp. 1-20). London: Springer

Weingaertner, D., Bellon, O. R. P., Silva, L., & Cat, M. N. (2008). Newborn's biometric identification: Can it be done? In *VISAPP (1)* (pp. 200–205). Retrieved from https://www.researchgate.net/profile/Luciano_Silva10/publication/221415455_Newborn’s_Biometric_Identification_Can_it_be_done/links/0c96052b8327747805000000.pdf

World’s Children.org. (n.d.). Prevent Trafficking. https://www.worldschildren.org/child-trafficking-statistics/. Retrieved May 29, 2020

Xia, X., & Ogorman, L. (2003). Innovations in fingerprint capture devices. *Pattern Recognition,36*(2), 361-369. doi:10.1016/s0031-3203(02)00036-5

Yager, N., & Dunstone, T. (2007). Worms, Chameleons, Phantoms and Doves: New Additions to the Biometric Menagerie. *2007 IEEE Workshop on Automatic Identification Advanced Technologies*. doi:10.1109/autoid.2007.380583

Yager, N., & Dunstone, T. (2010). The biometric menagerie. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*(2), 220–30.

Zhou, J., Chen, F., Wu, N., & Wu, C. (2009). Crease detection from fingerprint images and its applications in elderly people. *Pattern Recognition*, *42*(5), 896–906. doi: 10.1016/j.patcog.2008.09.011

Zibran, M.F. (2012). Biometric Authentication: The Security Issues. *University of Saskatchewan Technical Report*.

APPENDIX A

Table A.1 Subject Age Groups

SID	0-6 Months	7-12 Months	>12 Months
1	X	X	
2			X
3		X	X
4			X
5	X	X	
6		X	X
7		X	X
8	X	X	X
9	X	X	X
10		X	
11		X	X
12		X	X
13	X	X	X
14		X	X
15	X	X	X
16	X	X	
17	X	X	X
19	X	X	X
20	X		
21	X	X	X
24	X	X	X
25	X	X	
26		X	
27	X	X	X
28	X	X	X
29		X	X
30	X	X	X
31	X	X	X
34	X	X	X
36	X		
37	X	X	X
38	X	X	
39	X	X	X
40		X	X
41	X	X	
42	X	X	X
44	X	X	X
45	X		

46	X	X	X
47	X		X
48		X	X
49	X	X	X
50	X	X	
51	X		
52	X		
53		X	X
54	X		
55	X	X	X
56	X	X	
57		X	
63			X
64			X
67			X
68			X
72	X	X	
73			X
74			X
75			X
76			X
77			X
78			X
79			X
80			X
83	X	X	
84			X
85			X
87			X
88			X
89			X
90			X
91		X	X
92			X
93			X
94			X
98			X

APPENDIX B

Table B.1 Animal Classification Breakdown: 0-6 Months and 7-12 Months Overlap

Subject ID	Age Group (Months)	Animal Classification	Genuine Score	Impostor Score
LI1	0-6	Normal	43.00	36.80
LI1	7-12	Normal	44.28	35.23
RI1	0-6	Phantom	38.50	36.16
RI1	7-12	Normal	37.89	34.99
LI5	0-6	Normal	39.00	37.96
LI5	7-12	Normal	50.17	35.83
RI5	0-6	Normal	49.75	37.15
RI5	7-12	Normal	47.33	36.52
LI8	0-6	Normal	45.80	39.85
LI8	7-12	Normal	42.04	36.18
RI8	0-6	Normal	41.80	39.13
RI8	7-12	Worm	44.04	36.90
LI9	0-6	Dove	62.77	34.63
LI9	7-12	Dove	82.28	33.01
RI9	0-6	Normal	52.10	38.11
RI9	7-12	Dove	68.16	33.71
LI13	0-6	Normal	42.55	37.82
LI13	7-12	Normal	56.34	35.43
RI13	0-6	Normal	41.96	38.01
RI13	7-12	Worm	42.83	36.85
LI15	0-6	Normal	42.83	38.83
LI15	7-12	Normal	44.47	35.86
RI15	0-6	Normal	43.17	39.16
RI15	7-12	Normal	45.18	37.14
LI16	0-6	Normal	41.99	37.48
LI16	7-12	Normal	50.59	37.15
RI16	0-6	Normal	47.03	39.39
RI16	7-12	Normal	49.15	37.33
LI17	0-6	Normal	40.48	37.46
LI17	7-12	Phantom	37.07	34.25
RI17	0-6	Normal	37.57	36.90
RI17	7-12	Normal	49.64	35.56
LI19	0-6	Normal	58.43	37.55
LI19	7-12	Normal	46.31	36.49
RI19	0-6	Normal	45.40	37.09
RI19	7-12	Normal	58.26	34.61
LI21	0-6	Normal	52.03	37.60

LI21	7-12	Normal	60.07	33.77
RI21	0-6	Dove	60.59	36.67
RI21	7-12	Dove	71.96	33.59
LI24	0-6	Dove	96.27	34.90
LI24	7-12	Dove	73.36	29.47
RI24	0-6	Dove	64.46	35.68
RI24	7-12	Normal	51.79	34.88
LI25	0-6	Normal	43.33	40.17
LI25	7-12	Normal	46.70	34.74
RI25	0-6	Normal	47.70	39.23
RI25	7-12	Normal	39.31	35.35
LI27	0-6	Normal	47.00	40.19
LI27	7-12	Worm	43.37	37.31
RI27	0-6	Normal	44.35	39.33
RI27	7-12	Worm	44.25	38.31
LI28	0-6	Dove	92.00	28.21
LI28	7-12	Normal	59.38	32.78
RI28	0-6	Dove	111.33	29.09
RI28	7-12	Normal	51.49	32.67
LI30	0-6	Normal	40.17	37.22
LI30	7-12	Normal	45.94	36.14
RI30	0-6	Normal	58.05	37.88
RI30	7-12	Normal	48.92	35.93
LI31	0-6	Normal	47.35	35.38
LI31	7-12	Dove	81.54	30.16
RI31	0-6	Dove	66.15	34.34
RI31	7-12	Dove	66.77	32.38
LI34	0-6	Normal	44.72	38.75
LI34	7-12	Normal	45.19	36.69
RI34	0-6	Normal	40.78	37.32
RI34	7-12	Normal	45.68	35.75
LI37	0-6	Normal	44.00	39.69
LI37	7-12	Normal	45.12	37.13
RI37	0-6	Worm	41.67	40.09
RI37	7-12	Normal	45.38	35.90
LI38	0-6	Dove	60.15	35.31
LI38	7-12	Normal	90.67	36.34
RI38	0-6	Normal	52.11	37.16
RI38	7-12	Worm	44.17	37.01
LI39	0-6	Chameleon	74.33	41.22
LI39	7-12	Normal	49.88	34.90
RI39	0-6	Normal	45.83	38.84
RI39	7-12	Normal	54.42	35.12
LI41	0-6	Normal	49.06	41.15

LI41	7-12	Normal	44.38	36.21
RI41	0-6	Normal	45.29	40.28
RI41	7-12	Normal	50.79	35.57
LI42	0-6	Normal	45.55	38.42
LI42	7-12	Normal	46.73	36.20
RI42	0-6	Normal	52.32	39.21
RI42	7-12	Normal	45.94	37.55
LI44	0-6	Normal	63.50	38.38
LI44	7-12	Normal	48.33	34.50
RI44	0-6	Normal	49.00	39.23
RI44	7-12	Dove	60.69	33.70
LI46	0-6	Normal	92.00	37.35
LI46	7-12	Normal	69.36	34.71
RI46	0-6	Normal	42.50	38.33
RI46	7-12	Normal	51.22	34.86
LI49	0-6	Normal	41.74	38.01
LI49	7-12	Worm	38.00	36.86
RI49	0-6	Normal	40.28	38.26
RI49	7-12	Normal	47.33	35.47
LI50	0-6	Dove	64.78	35.24
LI50	7-12	Dove	152.67	26.84
RI50	0-6	Dove	61.25	32.85
RI50	7-12	Dove	86.67	29.20
LI55	0-6	Normal	44.57	39.24
LI55	7-12	Normal	95.67	35.03
RI55	0-6	Normal	46.38	39.83
RI55	7-12	Chameleon	88.00	37.51
LI56	0-6	Normal	37.55	37.20
LI56	7-12	Normal	68.97	35.93
RI56	0-6	Normal	42.00	38.43
RI56	7-12	Normal	45.55	36.36
LI72	0-6	Normal	42.13	38.14
LI72	7-12	Normal	43.50	35.74
RI72	0-6	Normal	42.83	39.34
RI72	7-12	Worm	41.10	36.73
LI83	0-6	Normal	43.00	37.34
LI83	7-12	Normal	47.25	35.99
RI83	0-6	Normal	43.00	35.89
RI83	7-12	Normal	45.00	35.43

Table B.2 Animal Classification Breakdown: 7-12 Months and >12 Months Overlap

Subject ID	Age Group (Months)	Animal Classification	Genuine Score	Impostor Score
LI3	7-12	Worm	37.67	37.41
LI3	>12	Dove	469.33	22.17
RI3	7-12	Normal	48.00	39.67
RI3	>12	Normal	84.00	24.77
LI6	7-12	Normal	48.17	38.19
LI6	>12	Worm	38.67	32.31
RI6	7-12	Normal	39.00	36.05
RI6	>12	Worm	45.17	33.46
LI7	7-12	Normal	41.74	36.35
LI7	>12	Normal	46.45	31.83
RI7	7-12	Normal	43.62	35.04
RI7	>12	Normal	45.27	31.40
LI8	7-12	Normal	42.04	35.97
LI8	>12	Normal	51.30	34.26
RI8	7-12	Worm	44.04	36.62
RI8	>12	Normal	47.27	30.99
LI9	7-12	Dove	82.28	32.96
LI9	>12	Dove	105.26	27.38
RI9	7-12	Normal	68.16	33.88
RI9	>12	Normal	87.00	29.28
LI11	7-12	Dove	145.30	29.75
LI11	>12	Dove	146.27	24.89
RI11	7-12	Dove	94.17	27.92
RI11	>12	Normal	97.77	26.33
LI12	7-12	Normal	46.50	36.37
LI12	>12	Normal	94.17	25.89
RI12	7-12	Normal	48.79	34.83
RI12	>12	Normal	76.90	26.16
LI13	7-12	Normal	56.34	35.08
LI13	>12	Normal	69.83	31.83
RI13	7-12	Normal	42.83	36.62
RI13	>12	Normal	59.50	32.65
LI14	7-12	Dove	74.93	32.02
LI14	>12	Dove	135.02	25.71
RI14	7-12	Dove	78.37	32.52
RI14	>12	Dove	119.18	26.38
LI15	7-12	Normal	44.47	35.65
LI15	>12	Normal	41.09	31.71
RI15	7-12	Worm	45.18	36.76
RI15	>12	Worm	40.50	31.92
LI17	7-12	Normal	37.07	34.12

LI17	>12	Normal	67.63	31.30
RI17	7-12	Normal	49.64	35.51
RI17	>12	Normal	136.00	31.50
LI19	7-12	Normal	46.31	36.08
LI19	>12	Normal	108.91	28.82
RI19	7-12	Normal	58.26	34.22
RI19	>12	Dove	155.28	26.75
LI21	7-12	Normal	60.07	33.61
LI21	>12	Normal	176.33	29.09
RI21	7-12	Dove	71.96	33.45
RI21	>12	Normal	123.17	30.42
LI24	7-12	Dove	73.36	29.32
LI24	>12	Dove	137.34	22.75
RI24	7-12	Normal	51.79	34.68
RI24	>12	Normal	76.01	28.91
LI27	7-12	Worm	43.37	37.00
LI27	>12	Normal	78.15	32.45
RI27	7-12	Worm	44.25	37.95
RI27	>12	Normal	133.17	30.55
LI28	7-12	Normal	59.38	32.55
LI28	>12	Normal	91.37	27.25
RI28	7-12	Normal	51.49	32.41
RI28	>12	Normal	97.63	29.66
LI29	7-12	Normal	61.22	34.88
LI29	>12	Normal	72.34	30.12
RI29	7-12	Normal	50.18	34.65
RI29	>12	Normal	103.66	31.69
LI30	7-12	Normal	45.94	35.79
LI30	>12	Worm	46.51	32.03
RI30	7-12	Normal	48.92	35.55
RI30	>12	Normal	47.55	31.55
LI31	7-12	Dove	81.54	29.82
LI31	>12	Dove	103.92	22.69
RI31	7-12	Normal	66.77	32.25
RI31	>12	Normal	58.91	24.31
LI34	7-12	Normal	45.19	36.38
LI34	>12	Normal	81.08	29.50
RI34	7-12	Normal	45.68	35.50
RI34	>12	Normal	90.00	29.35
LI37	7-12	Worm	45.12	36.76
LI37	>12	Normal	68.41	32.44
RI37	7-12	Normal	45.38	35.67
RI37	>12	Normal	114.08	29.50
LI39	7-12	Normal	49.88	34.76

LI39	>12	Normal	72.73	30.17
RI39	7-12	Normal	54.42	35.04
RI39	>12	Normal	99.03	27.75
LI40	7-12	Normal	79.30	33.68
LI40	>12	Normal	61.59	31.56
RI40	7-12	Dove	86.96	33.24
RI40	>12	Normal	59.16	32.05
LI42	7-12	Normal	46.73	35.84
LI42	>12	Normal	44.02	31.00
RI42	7-12	Normal	45.94	37.29
RI42	>12	Worm	47.43	32.94
LI44	7-12	Normal	48.33	34.47
LI44	>12	Normal	51.14	30.74
RI44	7-12	Normal	60.69	33.69
RI44	>12	Normal	63.71	31.16
LI46	7-12	Normal	69.36	34.46
LI46	>12	Normal	87.88	28.49
RI46	7-12	Normal	51.22	34.53
RI46	>12	Normal	69.61	29.05
LI48	7-12	Worm	44.09	36.63
LI48	>12	Normal	56.02	31.46
RI48	7-12	Normal	53.08	37.63
RI48	>12	Normal	53.07	30.58
LI49	7-12	Worm	38.00	36.59
LI49	>12	Worm	41.51	33.69
RI49	7-12	Normal	47.33	35.14
RI49	>12	Worm	42.34	34.10
LI53	7-12	Normal	63.57	35.56
LI53	>12	Normal	79.55	30.24
RI53	7-12	Normal	69.59	35.90
RI53	>12	Normal	61.57	31.45
LI55	7-12	Normal	95.67	34.69
LI55	>12	Normal	52.44	32.97
RI55	7-12	Chameleon	88.00	37.09
RI55	>12	Normal	63.64	31.89
LI91	7-12	Normal	45.22	33.05
LI91	>12	Normal	42.80	30.47
RI91	7-12	Dove	116.83	32.44
RI91	>12	Normal	41.77	29.62

Table B.3 Animal Classification Breakdown: Overlap Between All Three Groups

Subject ID	Age Group (Months)	Animal Classification	Genuine Score	Impostor Score
LI8	0-6	Normal	45.80	39.72
LI8	7-12	Normal	42.04	36.06
LI8	>12	Normal	51.30	34.78
RI8	0-6	Normal	41.80	39.11
RI8	7-12	Worm	44.04	36.77
RI8	>12	Normal	47.27	31.37
LI9	0-6	Dove	62.77	34.26
LI9	7-12	Dove	82.28	32.86
LI9	>12	Dove	105.26	27.36
RI9	0-6	Normal	52.10	38.05
RI9	7-12	Dove	68.16	33.68
RI9	>12	Normal	87.00	29.18
LI13	0-6	Normal	42.55	37.97
LI13	7-12	Normal	56.34	35.26
LI13	>12	Normal	69.83	32.63
RI13	0-6	Normal	41.96	38.14
RI13	7-12	Worm	42.83	36.77
RI13	>12	Normal	59.50	33.17
LI15	0-6	Normal	42.83	38.88
LI15	7-12	Normal	44.47	35.82
LI15	>12	Normal	41.09	32.21
RI15	0-6	Normal	43.17	39.19
RI15	7-12	Worm	45.18	37.03
RI15	>12	Normal	40.50	32.46
LI17	0-6	Normal	40.48	37.36
LI17	7-12	Normal	37.07	34.25
LI17	>12	Normal	67.63	31.42
RI17	0-6	Phantom	37.57	36.90
RI17	7-12	Normal	49.64	35.54
RI17	>12	Normal	136.00	31.55
LI19	0-6	Normal	58.43	37.47
LI19	7-12	Normal	46.31	36.33
LI19	>12	Normal	108.91	29.32
RI19	0-6	Normal	45.40	36.97
RI19	7-12	Normal	58.26	34.48
RI19	>12	Dove	155.28	27.18
LI21	0-6	Normal	52.03	37.74
LI21	7-12	Normal	60.07	33.78
LI21	>12	Normal	176.33	29.64
RI21	0-6	Normal	60.59	36.83

RI21	7-12	Dove	71.96	33.70
RI21	>12	Normal	123.17	30.77
LI24	0-6	Dove	96.27	34.65
LI24	7-12	Dove	73.36	29.40
LI24	>12	Dove	137.34	22.92
RI24	0-6	Dove	64.46	35.58
RI24	7-12	Normal	51.79	34.83
RI24	>12	Normal	76.01	29.14
LI27	0-6	Normal	47.00	40.15
LI27	7-12	Worm	43.37	37.24
LI27	>12	Normal	78.15	32.52
RI27	0-6	Normal	44.35	39.31
RI27	7-12	Worm	44.25	38.24
RI27	>12	Normal	133.17	30.88
LI28	0-6	Dove	92.00	27.85
LI28	7-12	Normal	59.38	32.70
LI28	>12	Normal	91.37	27.59
RI28	0-6	Dove	111.33	28.89
RI28	7-12	Normal	51.49	32.60
RI28	>12	Normal	97.63	29.80
LI30	0-6	Normal	40.17	37.32
LI30	7-12	Normal	45.94	36.06
LI30	>12	Normal	46.51	32.46
RI30	0-6	Normal	58.05	38.13
RI30	7-12	Normal	48.92	35.85
RI30	>12	Normal	47.55	32.10
LI31	0-6	Normal	47.35	35.36
LI31	7-12	Dove	81.54	29.99
LI31	>12	Normal	103.92	23.39
RI31	0-6	Dove	66.15	34.48
RI31	7-12	Dove	66.77	32.34
RI31	>12	Normal	58.91	24.70
LI34	0-6	Normal	44.72	38.76
LI34	7-12	Normal	45.19	36.59
LI34	>12	Normal	81.08	30.04
RI34	0-6	Normal	40.78	37.35
RI34	7-12	Normal	45.68	35.69
RI34	>12	Normal	90.00	29.87
LI37	0-6	Normal	44.00	39.64
LI37	7-12	Worm	45.12	36.93
LI37	>12	Normal	68.41	32.68
RI37	0-6	Worm	41.67	40.00
RI37	7-12	Normal	45.38	35.74
RI37	>12	Normal	114.08	29.80

LI39	0-6	Chameleon	74.33	40.81
LI39	7-12	Normal	49.88	34.83
LI39	>12	Normal	72.73	30.30
RI39	0-6	Normal	45.83	38.62
RI39	7-12	Normal	54.42	35.02
RI39	>12	Normal	99.03	27.63
LI42	0-6	Normal	45.55	38.51
LI42	7-12	Normal	46.73	36.12
LI42	>12	Normal	44.02	31.78
RI42	0-6	Normal	52.32	39.40
RI42	7-12	Normal	45.94	37.52
RI42	>12	Worm	47.43	33.61
LI44	0-6	Normal	63.50	38.40
LI44	7-12	Normal	48.33	34.51
LI44	>12	Normal	51.14	31.16
RI44	0-6	Normal	49.00	39.29
RI44	7-12	Normal	60.69	33.79
RI44	>12	Normal	63.71	31.63
LI46	0-6	Dove	92.00	36.96
LI46	7-12	Normal	69.36	34.63
LI46	>12	Normal	87.88	29.00
RI46	0-6	Normal	42.50	38.24
RI46	7-12	Normal	51.22	34.78
RI46	>12	Normal	69.61	29.51
LI49	0-6	Normal	41.74	38.11
LI49	7-12	Worm	38.00	36.77
LI49	>12	Worm	41.51	34.13
RI49	0-6	Normal	40.28	38.40
RI49	7-12	Normal	47.33	35.44
RI49	>12	Worm	42.34	34.43
LI55	0-6	Normal	44.57	39.39
LI55	7-12	Normal	95.67	35.03
LI55	>12	Normal	52.44	33.47
RI55	0-6	Normal	46.38	39.87
RI55	7-12	Chameleon	88.00	37.46
RI55	>12	Normal	63.64	32.62

APPENDIX C

Table C.1 Individual SSI Values for 0-6 Months and 7-12 Months Overlap

SID	SSI
LI1	0.004667
LI13	0.032361
LI15	0.007835
LI16	0.019901
LI17	0.010833
LI19	0.028122
LI21	0.020592
LI24	0.054404
LI25	0.014782
LI27	0.010707
LI28	0.076135
LI30	0.013584
LI31	0.079935
LI34	0.004876
LI37	0.006469
LI38	0.070559
LI39	0.058366
LI41	0.015736
LI42	0.005805
LI44	0.036185
LI46	0.052669
LI49	0.009035
LI5	0.026269
LI50	0.204028
LI55	0.118494
LI56	0.072668
LI72	0.006395
LI8	0.01215
LI83	0.010305
LI9	0.045248
RI1	0.003055
RI13	0.003353
RI15	0.006597
RI16	0.00685
RI17	0.028068
RI19	0.03027
RI21	0.027224
RI24	0.029345

RI25	0.021367
RI27	0.002372
RI28	0.138548
RI30	0.021582
RI31	0.004754
RI34	0.011886
RI37	0.012938
RI38	0.01837
RI39	0.021631
RI41	0.016725
RI42	0.015232
RI44	0.029882
RI46	0.021685
RI49	0.017534
RI5	0.005769
RI50	0.059343
RI55	0.096344
RI56	0.009498
RI72	0.007242
RI8	0.007321
RI83	0.004743
RI9	0.038484

Table C.2 Individual SSI Values for 7-12 Months and >12 Months Overlap

SID	SSI
LI11	0.011473
LI12	0.112799
LI13	0.032074
LI14	0.139654
LI15	0.012013
LI17	0.070941
LI19	0.145651
LI21	0.268901
LI24	0.148638
LI27	0.081064
LI28	0.074943
LI29	0.027963
LI3	0.998268
LI30	0.008797
LI31	0.054291
LI34	0.084472
LI37	0.054751
LI39	0.053863

LI40	0.04123
LI42	0.012826
LI44	0.010773
LI46	0.044969
LI48	0.030056
LI49	0.010522
LI53	0.038919
LI55	0.099972
LI6	0.025827
LI7	0.015095
LI8	0.021766
LI9	0.05467
LI91	0.008178
RI11	0.009099
RI12	0.067985
RI13	0.039596
RI14	0.095377
RI15	0.015563
RI17	0.199805
RI19	0.224887
RI21	0.118549
RI24	0.057533
RI27	0.206218
RI28	0.106829
RI29	0.123791
RI3	0.090048
RI30	0.00978
RI31	0.025817
RI34	0.103409
RI37	0.159417
RI39	0.104463
RI40	0.064312
RI42	0.010619
RI44	0.009095
RI46	0.044369
RI48	0.016284
RI49	0.011785
RI53	0.021184
RI55	0.057567
RI6	0.015462
RI7	0.009213
RI8	0.014978

Table C.3 Individual SSI Values for Overlap of All 3 Groups

SID	G1-G2	G2-G3	G1-G3
LI13	0.032493	0.03177	0.064263
LI15	0.008013	0.011443	0.015925
LI17	0.010662	0.070942	0.064245
LI19	0.028141	0.145583	0.118177
LI21	0.020734	0.268869	0.287904
LI24	0.054312	0.148617	0.098714
LI27	0.010759	0.081116	0.074124
LI28	0.076231	0.074877	0.001583
LI30	0.01367	0.00842	0.018476
LI31	0.07999	0.053935	0.133649
LI34	0.00513	0.084336	0.086427
LI37	0.006771	0.054722	0.05866
LI39	0.05817	0.053839	0.024589
LI42	0.006157	0.011811	0.015943
LI44	0.036186	0.010108	0.033112
LI46	0.052593	0.044733	0.02071
LI49	0.009172	0.01014	0.009201
LI55	0.118523	0.099958	0.022771
LI8	0.012122	0.02161	0.017086
LI9	0.045209	0.054621	0.0995
RI13	0.00376	0.039406	0.04213
RI15	0.006824	0.015114	0.016735
RI17	0.028076	0.199802	0.227823
RI19	0.030274	0.224858	0.254953
RI21	0.027261	0.118535	0.145299
RI24	0.029338	0.057493	0.030555
RI27	0.002495	0.206209	0.206201
RI28	0.138567	0.106835	0.031732
RI30	0.021757	0.009227	0.027987
RI31	0.005158	0.025325	0.028123
RI34	0.011958	0.103307	0.115058
RI37	0.01305	0.159372	0.169017
RI39	0.021528	0.104503	0.125549
RI42	0.015372	0.009678	0.017529
RI44	0.02987	0.008571	0.038333
RI46	0.021675	0.044225	0.065829
RI49	0.017688	0.011772	0.010347
RI55	0.096356	0.057403	0.043268
RI8	0.007494	0.014523	0.021892
RI9	0.038458	0.044774	0.083222