IDENTIFICATION, CLASSIFICATION AND MODELLING OF TRADITIONAL AFRICAN DANCES USING DEEP LEARNING TECHNIQUES

by

Adebunmi Odefunso

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Technology

West Lafayette, Indiana

May 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. James Mohler, Chair

Department of Computer Graphics and Technology

Dr. Victor Chen

Department of Computer Graphics and Technology

Dr. Esteban Garcia

Department of Computer Graphics and Technology

Dr. Cheryl Zhen Yu Qian

Department of Interaction Design and Industrial Design

Approved by:

Dr. Kathryne Newton

Dedicated to God and my late father Chief A.I Odefunso

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. James Mohler for his dedication, support and guidance throughout my PhD study and research. He gave me a roadmap at the beginning of my program which made this journey a memorable one. I would also thank all the members of my thesis committee, they contributed in a great measure to defining the path for my research. Dr. Esteban Garcia, Dr. Victor Chen and Dr. Zhen Yu Qian thanks a lot.

I also want to appreciate my supervisor, Mr. Don Brier and his wife Mrs. Cheryl Brier for their support and giving all, they can to make this academic program a success. I will not forget those babysitting days! All the professors I went through at Purdue were tremendous, some who I cannot but mention are Dr. Baijian Yang, Dr. John Springer, Dr. Tim McGraw, Dr. Vetria Byrd and Dr. Kristen Seward. Thank you all.

The Nigeria Students Association and African Christian Fellowship are two big families I belong to, I appreciate you all for your contributions in various ways. To the great elders, Prof. and Mrs. Ajuwon, Uncle Charles and Aunty Elizabeth, Aunty Betty, Aunty Anita, Mr. and Mrs. Ayano, Mrs. Amaka, Pastor and Mrs. Enefazu; thank you. Sis Opeadura, Sis Bidemi, Bro Femi, Bro Paul, Bro Emmanuel, Bro Kayode, Dr. Osho, Bro and Mrs. Ayo Aderibigbe: I am grateful to you all. Pastor and Mrs. Demola Dada, thank you for your support.

My appreciation goes to my siblings-Rev. Dr. and Mrs. Oyedepo, Mr. and Mrs. Kehinde Bolaji, Dr. and Dr. Tola Bolaji, Mr. and Mrs. Abodunrin, Mr. and Mrs. Adeyemo, Mrs. Oke, Mrs. Odeshina, Adebisi, Adebola, Adebanke, Dr. and Mrs. Adeoluwa, Mr. and Mrs. Folorunso and Adeniyi- thanks for your support. I owe my parents a lot. My mum (Mrs. Paulina Odefunso), Dn. and Mrs. Bolaji, Mrs. Adebimpe, Pastor and Mrs. Olugbenro, Uncle Monday and Aunty Glory, thank you all for your support in diverse ways and capacity.

Dr. and Dr. (Mrs.) Tayo Adedokun, The Jimos, The Adeagbos, Rev. G.OKehinde , Bro Sam, Alfa Ganiyu and Bro Ezekiel; thank you all for your support, guidance and prayers.

I want to thank my husband, Dr. Fola Bolaji, this would not have been possible without your support, provisions, prayers, and understanding. Thank you. Smith, Donald and Catherine, thanks a lot for helping mummy to achieve this.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Human action recognition continues to evolve and is examined better using deep learning techniques. Several successes have been recorded in the field of action recognition but only very few has focused on dance. This is because dance actions and, especially Traditional African dance, are long and involve fast movement of body parts. This research proposes a novel framework that applies data science algorithms to the field of cultural preservation by applying various deep learning techniques to identify, classify and model Traditional African dances from videos. Traditional African dances are important part of the African culture and heritage. Digital preservation of these dances in their myriad forms is a problem. The dance dataset was constituted using freely available YouTube videos. Three Traditional African dances – Adowa, Bata and Swange – were used for the dance classification process. Two Convolutional Neural Network (CNN) models were used for the classification and they achieved an accuracy of 97% and 98% respectively. Sound classification of Adowa, Bata and Swange drum ensembles were also carried out; an accuracy of 96% was achieved. Human Pose Estimation Algorithms were applied to the Sinte dance. A model of Sinte dance, which can be exported to other environments, was obtained.

CHAPTER 1. INTRODUCTION

1.1 Introduction

Cultural preservation encompasses preserving the tangible and intangible culture of a people or nation. This includes studying and documenting languages, preserving heritage sites in different countries, encouraging the use of indigenous languages and documenting how a group of people go about one thing or the other. When culture is not preserved, human diversity fades and important aspects of culture are lost. Traditional dances constitute a significant part of the cultural heritage around the world. This aspect of intangible cultural heritage needs to be preserved. There are many benefits to preserving cultural heritage (Tuan & Navrud, 2008). These benefits will be discussed later in chapter two.

The UNESCO Convention for the Safeguarding of the Intangible Cultural Heritage, 2003, defined intangible cultural heritage as "the practices, representations, expressions, knowledge, skills as well as the instruments, objects, artefacts and cultural spaces associated therewith – that communities, groups and, in some cases, individuals recognize as part of their cultural heritage" (page 97). The fragility of intangible cultures lies in their being sustained through lived circumstances and stored in human bodies and minds rather than in documents, artifacts, and forms of media.

Although dance is an important practice in most cultures of the world, it is a premier aspect of identity in the African cultures. The Traditional African dances, which usually involve rapid movement of various parts of the body at the same time, are dance variations that have been practiced over different generations. Such dance varies from one culture to another. In Nigeria, there are dances such as Bata, Ighogho, Swange to name a few. Other examples of Traditional African dance include the Adumu dance of the Masai tribe from Kenya, Kpalongo for Ghana, and Xhosa from South Africa. All these dances vary in the rate of body part movement and sequence. This study is based on contemporary theory of preservation that emphasizes the need to preserve cultural heritages and why they should be preserved.

Digital preservation can be defined as the preservation of materials in the digital format. This material might be digitized (that is, they were not digital but were made digital) or born digital (these are materials that were created originally in digital format). Legacy documents and artefacts that become digitized can be printed documents, pictures, photographs, handwritten documents, or physical objects. These materials are then digitized into images or digital documents using scanners, digital cameras, and some other technologies for preservation purposes and to make this material accessible to many people.

There are two traditional ways of preserving dance. The first involves a teacher teaching the dance techniques to the student or the older to younger generation. This involves a master dance teacher having a dance school where students or people who want to learn how to dance come to learn. This is common with most traditional cultures. As a matter of fact, this method is still used today. The other method is by documenting the dance techniques using media formats such as film, motion-capture, video and notation, etc.

Technology as a tool for preserving culture is increasingly growing. The application of computer software and technology in general focuses mainly on improving digitization and documentation of artifacts and sites as well as exploration and reconstruction of monuments (Pavlidis, Koutsoudis, Arnaoutoglou, Tsioukas, & Chamzas, 2007). The use of technology in cultural heritage preservation has become popular with the application of technology to every field. The field of tangible heritage preservation was quick to experience technology integration. The use of technology by computer graphics specialists has led to impressive results in preserving cultural heritage sites (Berndt & Carlos, 2000). Application of technology has been mostly the development of 3D models of artefacts and 3D reconstruction projects. The use of virtual reality is a recent innovation. The application of 3D modelling techniques is often used for the reconstruction of physical and natural heritage sites. It is a common place thing to see miniature models of heritage buildings like museums, historical buildings, and sites that is held in high esteem as part of the story of a community. The development of 3D modelling paved the way for the application of virtual reality in cultural preservation. Not only are models of these heritages constructed, virtual reality paved way for the construction of virtual heritages.

Automatic human action recognition is a complicated problem with which computer vision scientists are still grappling. This is because it involves categorizing and mining patterns of human poses from videos (Kishore, Kumar & Kumar, 2018). Human action is defined as a temporal variation of the human body. Many classification and mining activities have been carried out on a wide variety of datasets, but only few have been on dance classification and much less has been targeted towards the generation of data that can be used for cultural preservation. The complexity of most African dances makes it a challenging endeavor within the general field of activity recognition (Kapsouras, Karanikolos, Nikolaidis, & Tefas, 2013).

An example of automated human action recognition was a study conducted by Sangeeta, Anwaya, Manish and Jyoti (2012). The study generated automatic pose stick figure representation of Bharatanatyam (BN), an ancient Indian Classical Dance from motion capture data collected from a professional dancer.

1.2 Problem statement

A significant result of colonialism and globalism is cultural assimilation. The world is gradually becoming a village. As well-meaning as this may be, several traditional art forms, means of expressions, cultural identities, festivals and activities long practiced are being lost. Distinctive cultural heritage such as dance shows cultural diversities among people, and it is the cultural root for such societies. Technology has been used for preservation of several tangible cultural heritages such as monuments, shrines, tombs etc. However, less is being done with technology for intangible cultural preservation. Deep learning techniques have been used for the identification, classification and generation of human movement actions; however, it is not known if it can be used for dance cultural preservation, hence this study.

1.3 Objectives

The aims of this study are to:

a. Generate image and video data of Traditional African dances that can be used for action recognition and prediction studies as well as archived for cultural preservation purposes.
b. Develop a deep learning algorithm for the identification and classification of Traditional African dance.
c. Develop a deep learning algorithm for the modelling of a pose stick dance model that can be used for dance preservation.
d. Develop a framework for intangible cultural preservation using deep learning techniques.

1.4 Significance of the study

The findings of this study will provide an algorithm that can identify and classify Traditional African dance. It will also provide a digitized version of the dance and a pose-stick dance model that can be used for dance preservation. The study will also provide a Traditional African dance dataset for research in computer vison and artificial intelligence that is not readily available now. This study is justified in that the cultures of the African continent are becoming an endangered species, which makes it a state of emergency to investigate it using the available digital tools for their preservation and promotion. It is important that the African stories and culture be represented for them to be preserved.

1.5 Research Questions

This study would attempt to answer following research questions:

a. Is YouTube a sufficient source for generating Traditional African dance dataset (TAD^2)?
b. Are deep learning techniques suitable for intangible cultural preservation?
c. What deep learning technique(s) will identify and classify Traditional African dance poses?
d. Can deep learning techniques generate a 'tangible' dance model of Traditional African dance from videos?

1.6 Hypothesis

The following hypothesis would be tested:

a. YouTube videos will be enough to generate hundred thousand images for dance classification study.
b. Convolutional Neural Network (CNN) will classify Traditional African dance at 95% accuracy.
c. Deep learning techniques will generate a Traditional African dance model from videos at 95% accuracy.

1.7 Assumptions

The following assumptions are made for this study.

a. Traditional African dance in its various forms have similarities that will make dance model algorithm for one dance type suitable for another.
b. Traditional African dance have distinct characteristics that will make classification with deep learning possible.

1.8 Delimitations

The delimitations of this study include:

a. Traditional African dance from minority groups will be excluded in this study.
b. The dance classification study will exclude choreographic performances.

1.9 Limitations

The limitations of this project include:

a. The study is limited to Traditional African dances only.
b. All the images generated for the classification process belong to only three classes – Adowa, Bata and Swange. This decision is based on the number of video available per class. Much more videos were available for these three classes than others.
c. Another limitation is that pose stick generation from pose estimation algorithms was limited to only one dance class. The decision to use the dance class was because the costume used in the video did not occlude dancer's lower joints which was a major limitation with the other dance classes.
d. Another limitation is that the pose stick generated was that of a single dancer.

1.10 Definition of Terms

Culture preservation: It is essentially the documenting, preserving, and restoring languages, historic artefacts that are significant to a people, historical sites that preserve the history of a people and encouraging these acts among people. Cultural preservation is important for preserving human identities. Cultural preservation includes documenting and studying languages; preserving and restoring historic relics significant to a culture or heritage; and encouraging the preservation and use of indigenous or tribal languages and rituals.(https://www.daytranslations.com/blog/2017/09/cultural-preservation-9737/).

Intangible Cultural Heritage: According to UNESCO (2003) intangible cultural heritages are 'the practices, representations, expressions, knowledge, skills – as well as the instruments, objects, artefacts and cultural spaces associated therewith – that communities, groups and, in some cases, individuals recognize as part of their Cultural Heritage' . Examples include oral traditions, dance etc.

Deep Learning: This is a subset of the machine learning field that is modelled after how the human brain works which is termed as Artificial Intelligence. Deep learning connects neurons in layers like the brain which are referred to as Artificial Neural Networks.

Digital Preservation: Digital Preservation Coalition (2006) defined digital preservation as "all activities that are required to maintain access to digital materials beyond the limits of media failure or technological change. Those materials may be digital records created during the day-today business of an organization, i.e. "born-digital" materials created for a specific purpose (e.g. teaching resources), or the products of digitization projects"

Traditional Africa dances: These are dances which are indigenous to a culture and is part of the identity of a people. These dances are dances performed by the African ethnic groups which has been in existence almost as far back as the existence of the tribe.

1.11 Summary

The problem of Traditional African dance preservation and the need to preserve these dances has been identified. The objectives of the study includes the development of a data set that can be used for Traditional African dances study; identify, classify and model Traditional African dances using deep learning techniques was also discussed. The research questions, hypothesis, limitations, assumptions, and delimitations of the study were also discussed. Finally, some terms that might lead to confusions were also defined as they are intended for this study.

CHAPTER 2. REVIEW OF LITERATURE

2.1 Introduction

There are different parts, views, and opinions in relation to Traditional African Dances. In this chapter, the review of literature is focused on intangible cultural heritage, theory of cultural preservation, deep learning, image classification, dance classification, acting recognition algorithms and action generation algorithms were investigated.

2.2 Theoretical Framework

This study is based on the self-identity theory and the theory of historic preservation. Both theories formed the foundation on which the idea of using modern technology for dance preservation is based.

Self-identity theory draws attention to the notion of self and identity (Horowitz, 2012). The identity that a person has of him/herself determines how they will behave and what they will do. A disconnect from proper identity can lead to negative reactions and behaviors. The identity of an individual is engraved within a culture. Human beings have their conscious and unconscious parts, both of which provides information within that person. Social views in the community also provide external information to the person concerned. When both of this comes together, the interpretation given by the individual determines how the person will self-identify. Who a person will identify him/herself has is dependent on the social views he/she is surrounded with, as well as his/her unconscious parts? The culture one is born into or grow around is important to the identity of the individual. It is important that individuals have a proper self-identity else they are accountable to nobody however their life is lived. Relationship conflicts are inherent in self-identity is perverted.

The theory of historic preservation is focused on the maintenance, reconstruction, and the sustenance of sites, items, scenes, monuments and other important relics of great implication in the lives of the people. This theory emphasizes physical buildings and materials as important to the historical narratives of a people (Rodrigues, 1998). The efforts to preserve historic site has been led in many terms by developed counties. This is being spread to the developing countries as well. Preservation efforts can be seen around the world as monuments, shrines, parks, antiquities, graveyards etc. are being protected from going into extinction. These efforts have paid off in contributing to advancement in the field of tourism and recreation. It has protected some narratives which would have gone with the wind were the sites not preserved.

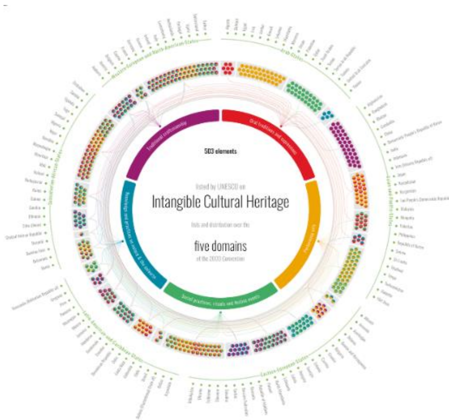
This study proposes the theory of intangible heritage promotion and preservation with artificial intelligence. The emphasis of this study is to develop a framework for intangible cultural heritage preservation using artificial intelligence. With the technical advancement of the 21st century where machines are being designed and used for achieving task completion at a speed that is not comparable to that of humans, it is important that this technology is utilized for a field whose elements seems to be uncountable, although important. The elements categorized as intangible cultural heritage are as numerous as the various kinds of people who live indifferent parts of the world. It is important that these elements are known and preserved as they serve important roles in the identity of people.

2.3 Intangible Cultural Heritage

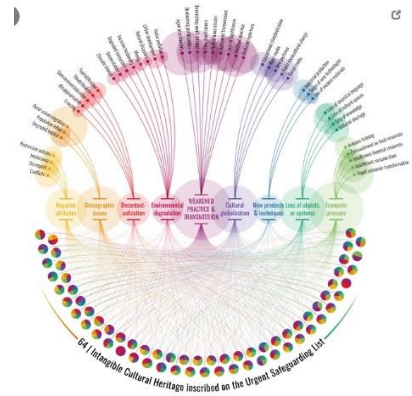
Intangible cultural heritage was defined in an article 2 of the UNESCO Convention for the Safeguarding of Intangible Cultural Heritage (CSICH) as: 'the practices, representations, expressions, knowledge, skills – as well as the instruments, objects, artefacts and cultural spaces associated therewith – that communities, groups and, in some cases, individuals recognize as part of their cultural heritage. This intangible cultural heritage, transmitted from generation to generation, is constantly recreated by communities and groups in response to their environment, their interaction with nature and their history, and provides them with a sense of identity and continuity, thus promoting respect for cultural diversity and human creativity'. This definition is important to the understanding of intangible cultural heritage. Examples of intangible cultural heritages include folklores, dance, cloth weaving, pottery, oral recitations, festivals etc.

The preservation of intangible cultural heritage was neglected for a long time because it does not appear as if it had any threats. However oral cultures, traditions, festivals which are important to the cultural identity, creativity and conservation of cultural diversity is gradually been lost to globalization, rural exodus, migration, knowledge scarcity and loss of practitioners.

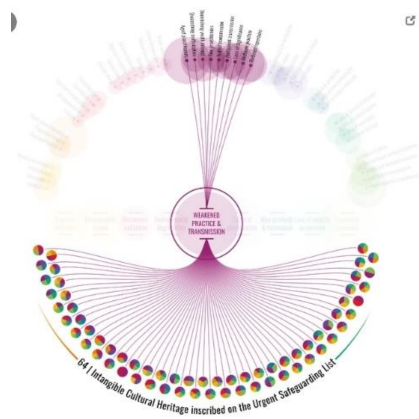
In the effort to preserve intangible cultural heritages, UNESCO has taken the lead in identify some cultural elements and making efforts to incrementally fund and take steps to preserve them. Currently, UNESCO is focused on Five hundred and forty-nine (549) elements of culture in one hundred and twenty-seven (127) countries, which is an improvement on the Five hundred and three (503) that were first listed in 2003 convention. These intangible heritages were divided into five major domains: oral traditions and expressions, social practices, rituals and festive events, knowledge and practices on nature and the universe, traditional craftsmanship and performing arts (figure 2.1a). The studies by UNESCO identified all the factors that are affecting intangible cultural heritage. Forty-six (46) factors which are categorized into nine (9) major threats (figure 2.1b). The major threats facing intangible cultural heritages are negative attitudes, demographic issues, de-contextualization, environmental degradation, weakened practice and transmission, cultural globalization, new products and techniques, loss of objects or system and economic pressure. Of all these pressures a major threat to all the cultural elements that UNESCO is safeguarding from going into extinction is weakened practice and transmission (figure 2.1c). This challenge has posed a major threat because of the means of transmitting cultural practices from one generation to another. The major means of intangible heritage transfer has been passing the tradition by practitioners from one generation to another. Due to globalization, some of these traditions are now practiced only weakly and, in some cases, the upcoming generation has lost interest in them. It is important that intangible cultural heritages are saved from going into extinction



(a)



(b)



(c)

Figure 2.1: a. Elements of intangible cultural heritage by UNESCO. (b) Threats and (c) Major threat to all. Source:ich.unesco.org

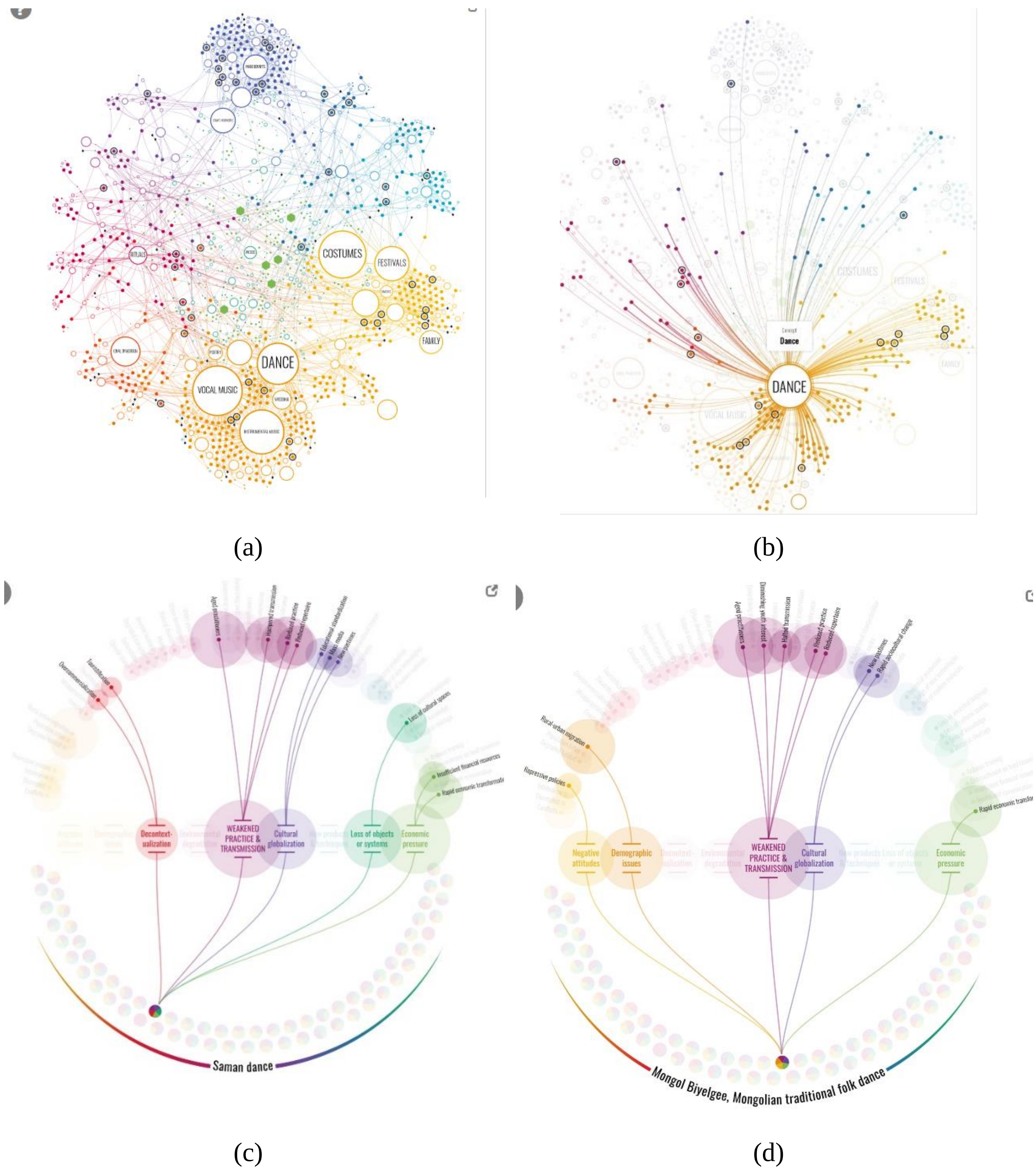
2.4 Traditional African Dance

Traditional African dances like those of other parts of the world are performed at different occasions. In Africa, dance does not serve the role of entertainment only, but rather dance tells a story. Dance is used to narrate history, transfer emotions, celebrate rites of passage, and helps to unite communities. The term Traditional African dance has been limiting because there are several dance types around the continent. The term “African dance” is usually associated with dances originating from the sub-Saharan and Western part of the continent. A major characteristic of African dances is that they are polycentric in nature.

Polycentricism in African dances means that different parts of the dancer’s body move to different rhythms in the music. These polycentric nature makes these dance styles quite difficult to master. Although this is a general feature, there several other characteristics that differentiate one dance from another. Another way to categorize Traditional African dances apart from movement is the occasion where those dances are performed. There are ceremonial dances e.g., Adowa performed in Ghana, ritual dances e.g., Kakilambe dance performed in Guinea, festival dance e.g., Bata performed in Nigeria, etc.

Dance has been identified as important cultural elements to be preserved. The list of the intangible heritage to be safeguarded by UNESCO is made up of several Traditional dances as presented in figure 2.2a. The size of the bubble shows the number of elements being preserved that are dances. Dance is not an isolated tradition, it is associated with other factors such a s music, musical instruments, poetry, religion etc. (figure 2.2b).

Traditional African dance is a fundamental component of Africa's social legacy, giving an imperative articulation of the rich philosophy of the continent, and the living memory of its social riches and its development throughout the different generations. The list of Traditional African dance is inexhaustible. It is to be noted that some dance styles are shared by different regions. The following are examples of tradition African dance: Agbeko, Agahu, Mohobelo, Unteyo, Jerusarema, Yankadi, Moribayasa, Agbekor, Macru, Kpanlogo, Adumu, Indlamu, Kete, Owaro, Sabar, Sunu, makossa, Entogoro, Akogo, Atilogwu, San, Aduma, Ekista, Agbekor, Mbira, Maasai, Batwa, Dogon, Ewegh, Laban, Mbira, Igogo, Igbin, Bata, Ohogho, Koroso, Swange, Ikpirikpi-ogu etc.



(a) (b)

(c) (d)

Figure 2.2: a. Dance as an important intangible heritage. (b). Dance is not an isolated cultural element, (c) and (d) threats to dance as a cultural heritage. Source:ich.unesco.org

2.4.1 Types

There are different school of thoughts to the way to categorize Traditional African dances; some people focus on the occasion while others focus on the body parts emphasized during dances. The major types of Traditional African dances are as follows:

a. Ritual dances: These dances are those which are part of a religious ritual. They are performed at special occasions to communicate and facilitate expressions of the people to the god of their land. These dances affirm the belief system of the people and the dance to be performed varies with what is being performed. Examples of this dance includes the Mbira dance of Zimbabwe, Igogo and Bata dance of Nigeria etc.
b. Ceremonial dances: These are as common as the ritual dances but are more performed because of the flexibility of the occasions where they are used. There is a dance type for most occasions ranging from childbirth, to marriage and even visitors welcome in different cultures. Dances of welcome are performed when a visitor is in town to express pleasure and respect towards the visitor. This is not only performed when the visitor is far from the culture, it can be as simple as the community leader visiting a school. The students gather to sing songs of praise to the visitor and express their respect and pleasure at having them around. Dances performed at different occasions such as marriage ceremony, graduation, child christening etc. also belongs to this category. Royal dances are performed for chiefs, kings, and important dignitaries in the community. This dances usually express the majesty of the dignitary that is present. Dances which are performed when youths become adults usually referred to as rites of passage and coming of age are also ceremonial dances. These dances are performed to showcase and celebrates people who have recently become adults in the community.
c. Griotic Dances: These dances are performed to tell a story. They are used to pass traditions and stories down from one generation to another. Sometimes they are used to tell the younger generation stories that emphasizes their values. The music that this dance is performed to usually have metaphorical statements that pass the culture of the people from one generation to another.
d. Communal dances: This is the heartbeat of the African communities that is losing its rhythm very fast. The life of the community is expressed in the gathering of the people. When people come together usually in small groups in their community, they perform

dances that gives them a sense of togetherness. This dance performances are becoming things of the past in most communities, especially cities. When the drumbeats sound, people start coming together. It is a time to connect with others and share the moment. It is usually fun for both old and young. These dances reinforce the community. Dances can be performed with the completion of a new project e.g. building of a town hall, seasonal changes, harvest, etc.

2.4.2 Importance of Traditional African dance preservation

According to Ilado (2017), 'Ivorian choreographer Alphonse Tierou explains in his book *The Eternal Law of African Dance* (1992) that Traditional dance is an essential element of Africa's cultural heritage. It has more power than gesture, more eloquence than word, more richness than writing, and because it expresses the most profound experiences of human beings, dance is a complete and self-sufficient language. It is the expression of life and of its permanent emotions of joy, love, sadness, hope, and without emotion there is no African dance. Traditional African dance is the soul of the community, it is a coupling power which can't be trivialized. Traditional African dances are important as a cultural heritage for these reasons.

Identity

Identity is a social classification characterized by trademark properties or practices that characterize what one's identity is. This identity is normally perceived by oneself as well as other people. In any case, an identity might be additionally ascribed to an individual or a group of people. Dance, a marker and image of character, is ground-breaking in making, strengthen, and evolving personality. How we realize what our identity is, who we ought to be and how (perspectives, acts, and feeling) are established in the mind, the pictures we see, and learning instructional method. The mind deciphers the pictures regarding information an individual has and embeds this in the memory stockpiling portions of the cerebrum as indicated by neuroscience (Hanna, 2015). Dance develops, fortifies, and deconstructs or questions conventional and evolving identity.

Individuals have numerous characters that can be communicated in the equivalent or in various moves. There might be a distinction between artists' demeanors of character and crowd recognitions. Congruity and change can happen after some time and at the same time. During pre-colonial and colonial periods, Westerners' impression of the personality of "African dance" was pessimistic. Dance was related with pagan religion. Christianity dismissed things animalistic, and the presentation of Christianity annihilated the raison d'etre of most of the dance types in Africa. Some portion of indigenous conviction frameworks, moves were prohibited at minister schools. Because of this most Africans needed to be shown how to play out an African dance move since they have been discarded by some misinformed parents. These children cease to cultivate the African values because their identity has been altered.

Europeans passed mainstream moral decisions: Within the Victorian casing of ethical quality, African moving was marked as indecent, brutal presentation. Not many Europeans acknowledged, for instance, that moves including a lot of pelvic movements could in addition to other things be a glorification of fertility identified with the desire for plentiful harvests, a worldview of life power, and a confirmation of life itself. Checked against Western theater workmanship expressive dance, African moves were viewed as crude and fascinating. Shaking of waist and hips vigorously does not connote any negativity in the African minds.

With independence, most African dances got some positive outlook, but some dance types never became popular again. Numerous African gatherings affirmed their national and different identities through dance. With a dance performance I can tell who is from one tribe or another. Most Traditional African dances differ in costumes. While some dances require the wearing of masks, some do not. According to Hanna (2019), the use of the name "African Dance" overlooks the multiplicity of the different identities that make up Africa. According to Mabingo (2012), the word 'African dance' should not be used again but rather 'African dances'. As of now, the impacts are inserted in Western also, African contemporary move studios, organizations, also, associations. Africa is treated as a nation. Dance(rs) from Africa are treated as elements without idea, hypothesis, theory, setting. It is not uncommon to find artists and students who don't know much about Africa learn and perform what they see as "West African dance" and also, gain a 'west Africa dance' experience. In this occurrence, the lavishly differing moves from West Africa are decreased to insignificant movements and wrappers.

Community and support

The community life in is no doubt a thing of excellence. Dance and music play a significant role in this. Dance bridges gaps in the society. Most African dances are performed during occasions. It communicates with people and bring several people together. Solo dancing is not as common as group dancing. A baby is welcome to the world with ceremonies and this is carried on until such exit this world. Dance in the African society is performed at births, marriage, funerals etc. 'Community' is a multivalent idea, subject to what a person perceives as important and derives pleasure in. Johnson (2016) a dance artist, detailed the experience of the term ‘Community’ that she experienced when she took a dance class in her thesis.' The class was a ‘West African' dance class in Philadelphia — assigned as a 'network based' class by the instructor of the class. The class, one of a few offered all through the city, is situated in West Philadelphia. It is an intergenerational class gone to by an assorted segment of members (race/ethnicity, sexual orientation, calling, class, age, capacity, and so on.) with a variety of inspirations and objectives for partaking in class (as made apparent through discussions and meetings). All are free to join in, paying little heed to past understanding or ability level in 'West African' dance. The beauty of the experience that she documented is in the fact that whether in the continent or outside it, most

African dance binds people together beyond partial participation. This binding force should not be allowed to go down the generations because of globalization and misconceptions. African dance is completely unique in relation to the Western style of dance. Customary African dance is a community activity, rather than a two ‘partner’ dance in male-female sets. Rather, the greater part of the dances are bunch exhibitions isolated by sexual orientation. The men dance for the ladies and the other way around, with all ages blending or having their own dance. This strengthens the ancestral jobs, both regarding the genders and furthermore as far as a gathering character.

One of the most striking pieces of customary African dance is the nature of the movement. African artists regularly can seclude specific pieces of their body and dance them to various pieces of the rhythm, with a few unique beats going on at the same time in the artist's body. This is joined by bigger movements, for example, kicks, jumps, and wide and fast swings of the arms. There are a wide range of explanations behind the different dances, all mirroring a piece of life. This can be a simple work tune to help make regular undertakings, for example, washing or tending fields, increasingly charming, yet the more complex dances are typically performed in view of some reason.

Music, tune, and dance are found in all community exercises, to transfer messages of congratulations, welcome, reactions, to reflect articulations of distress and compassion, in festivities, and in marriage arrangements and festivities, customs, births, passings, transitional experiences, hunting, and even political exercises.

Wellness

Traditional African dance promotes wellness. While some dance as tagged as healing dance due to religion, most dance promotes good mental health because of the support and relaxation it provides. Some dances in the African context exercise the whole body. The rapid movements are very good for the muscles. Aside from physical exercises, the mind is placed in a relaxed mood. There are some African dances categorized as war dances. These dance types build courage into the warriors that are going to the battlefield. Even when one person is failing in courage, the energy that comes from all the members of the group soon transfers courage into the weak warrior and he gains his courage back in a little time. Dancing reduces stress. Some African dances are called work dance. These dance moves are performed when people are at work. It is not uncommon to see women pounding yam, singing, and dancing in the kitchen. While this act can be termed barbaric by some, what it does to the African woman who is cooking is that she is not feeling stress as much as if she is just doing the cooking. Farming, making ridges, building of houses etc. have dances that are performed in relation to them. Songs that are song would be one that talks about the harvest that is coming after the action of cultivation. All of this, I presume are part of what made the older African population void of mental health issues and critical health concerns that are prevalent today.

There are some dances that are religiously performed to receive things from God. Some of these are dance performed by women who have infertility issues. While medical science has overtaken these, women have been blessed by children before by dancing in front of the Nigerian Osun Goddess. Rann (2015) documented the experiences of students in soul healing from their participating in Ghana dance classes in the University of California, in her PhD thesis.

African moves are among the most seasoned move conventions in presence; their structure is remarkably extraordinary on the grounds that the movements are indistinguishable from the music that administers the movements. The music is related with the communicated language of the individuals, which makes it practically unimaginable for outsiders to fathom the music of various African nations. In Africa there is no move that isn't joined by some type of music from the voice to symphonies of various percussive instruments. For quite a long time the move/music of African individuals has been passed between ages by a mouth to ear process Any society that is entirely dependent upon oral communication to transfer their culture between one generation to another is doomed to failure on account of the breakdown of the human memory and outside translation. The most ideal approach to amend this predicament is to give composed documentation to these moves. Since the moves are indistinguishable from the music. This led to the development of a framework called Greenotation, after Green Doris who developed it (Green, 2018). Right now, just can the music and moves all through Africa can be protected, and given interminability, yet additionally exhaustive proposition and thesis would now be able to be composed though beforehand, this couldn't be cultivated on the grounds that African move/music did not have a composed configuration.

2.5 Intangible Cultural Heritage Preservation

2.5.1 The neglect of intangible cultural heritage preservation

There have been little efforts directed at preserving intangible heritage globally before the UNESCO convention of 2003. The nature of intangible heritage is complex especially because of its interdependence on the tangible cultural heritage. Museums, for example Mystic Seaport in Connecticut (Bruggerman, 2009) have been used to preserve Traditional craftworks. But these efforts have not been enough for all the numerous intangible cultural heritages that exists. Most of the preservation efforts aimed at intangible cultural heritage can be accurately described as the documentation of intangible heritage.

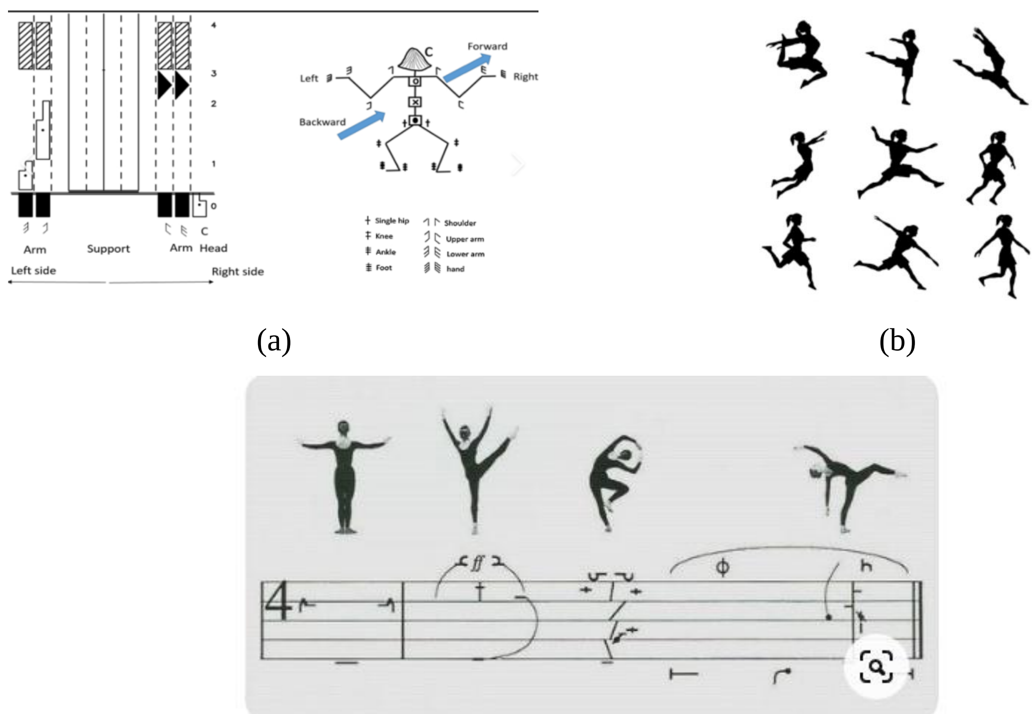
There is a significant gap between documenting and preserving culture. For example, there is a difference between documenting a language and preserving the language (Kroskrity, 2002). While documentary practices go a long way in contributing to preservation, they are not preservation in themselves. A fundamental truth about intangible cultural heritage is that they are practices. Thus, their preservation efforts must include establishing conditions that will communicate those cultural knowledge and practices from one generation to another. In this age, intangible cultural heritage digitization to make it easily accessible to this generation plays a vital role in preserving this culture. This study intends to focus on dance as an example of intangible heritage and use the emerging technology of artificial intelligence for the preservation efforts.

2.5.2 Dance Preservation

It is a popular in the dance industry that dance as well as digital media themselves are not easy to preserve because of their ephemeral natures; however Traditional dances have lasted a while to show that they could be preserved. While there have been many alterations to the different dance forms from one ancient culture to another, the bottom-line movement, poses and body shape are still largely maintained in those cultures.

Traditional dances are usually preserved by transmission. A master dance teacher having a dance school where students or people who wants to learn how to dance come to. This is common with most traditional cultures. As a matter of fact, this method is still be used today. The other method is by documenting the dance techniques using media formats.

Dance notation involves the use of symbols written as text, the approach for the interpretation of movement differs from one person to another. Three widely accepted notations used for Western dance are Labanotation, Laban Movement Analysis (LMA), and Benesh Movement Notation (BMN) (Goodridge, 1999).



(a) (b)

(c)

Figure 2.3 (a) Example of Labanotation score and corresponding body parts, (b) Example of Laban movement analysis,(c) Example of Benesh Movement Notation.
Sources a: Ikeuchi, Ma, Yan, Kudoh and Nakamura, 2018).b: retrieved from https://www.aromatherapyandmassage.com/laban-movement-analysis.html; c: retrieved from https://www.pinterest.com/pin/57069120258896733/

Dance Heritage Coalition (DHC) started digital techniques for preserving dance since the year2000. They digitize dance collections and provide financial aids to dance companies and other organizations interested in. Whereas organizations like Merce Cunningham Foundation have the infrastructure and funds that can be used to preserve choreographic works, most individual choreographers do not have that luxury at their disposal. Because of this, most dancers and choreographers' resort to YouTube, Pinterest, and other social platforms to keep, publish and store their works.

In 2002, the DHC started a three-stage venture, facing the Magnetic Media Crisis. Stage I centers around tape helplessness and organization out of date quality by making a database library, the National Dance Heritage Videotape Registry, to recognize those tapes and assortments that are in prompt need of safeguarding. Stage II will give unobtrusive subsidizing to re-acing, duplicating, or protection of jeopardized tapes chose from the Registry. For more data on this task or to round out a Registry poll if it is not too much trouble contact the DHC. To address issues concerning the exchange of simple tape to advance for safeguarding purposes, the DHC has started the Digital Video Preservation Reformatting Project.

Merce Cunningham Dance Company (MCDC) was an initiative developed by a former dancer Martha Graham to preserve the works of the choreographer Merce Cunningham, who died in 2009. The capsule includes sound recordings, costumes, dance script, and costume design etc. Motion-capture innovation has been effectively adjusted for media move by choreographers, for example, Merce's Hand-drawn spaces is a method for expanding choreographic prospects (Birringer, 2002). However, this method was not as effective because of the significant expense and concentrated hardware that is needed to keep these materials are not available. Another problem with having a lot of documentation is information retrieval.

If information cannot be easily retrieved, then the goal of the digitization has been forfeited. What Choreosave did to address these issues was the creation of a short-term conservation community-based digital repository called ChoreoSave. This repository utilizes an open source programming. Rather than storing numerous variants of dance films, a standard type of filing dance works was created. This plan protects the dance films and related components. This methodology challenges customary thoughts of digitizing dance using data association standards. By framing a standard submission template with customization options for users, the appraisal, choice, and ingestion procedures will turn into the clients' obligation, in this way streamlining the progression of substance from creator to storage. A computerized archival system proportional to MCDC is Siobhan Davies Dances Archives. It was the outcome of Sarah Whatley's doctoral investigations. It was seen as “the UK's first computerized move chronicle (http://www.coventry.ac.uk/cu/d/162/a/468, n.d.)."

The reason for ChoreoSave is to give a short-term protection administration to developing choreographers who are not familiar with the document storage in the library. It was borne out of the need to share dance techniques using video. The framework was guided by documented practices, advanced conservation, web-based social networking, and past strategies for saving move. Or maybe, it is intended to be an advanced curation instrument for safeguarding move movement, not execution. Figure 3 presents the potential components of a dance work that needs to be preserved.

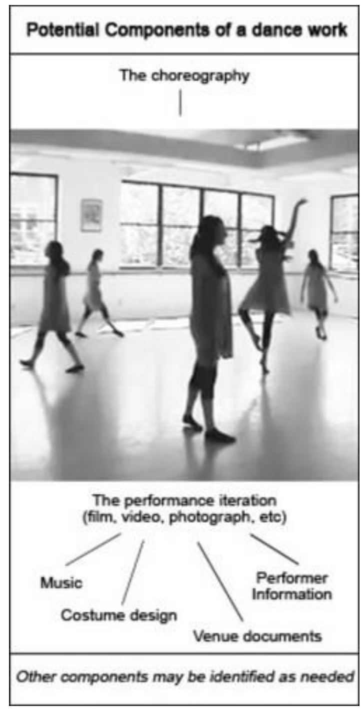


Figure 2.4: Choreosave’ s contents to be digitized. (Kim, 2012)

The study elucidated information concerning the development of the archival system by using questionnaires. A 10-question study was sent to every association involved. These questions range from the criteria for qualification, the criteria a media that will be accepted must meet and so on. Open-ended questions were used with the goal that remarkable bits of knowledge into dance preservation could be caught. Respondents could make their choice. The main inquiry of Survey A posed to move organizations to list the sorts of things that are acknowledged that can be used in preserving dance. There were 7 respondents who returned an aggregate of 23 reactions. As found in table 1, "Films/Videos”, “Photographs,” and “Printed Materials” were the most frequently listed items Table 1 and 2 are summaries of the information derived from the questionnaires.

Table 2.1. Items Used to Preserve Dance Works

Name of Item	Frequency1	Percentage of Total Responses%	Percentage of Total Respondents
Multi-faceted Approach	1	4%	14%
Music Scores and Media	1	4%	14%
Dance Scores	2	9%	29%
Lighting Cues and Designs	2	9%	29%
Production Information	2	9%	29%
Costume Sketches and Fabric	3	13%	43%
Films/Videos	4	17%	57%
Photographs	4	17%	57%
Printed Materials	4	17%	57%

Note. Results are from the findings of Kim, 2012.

Information about the items required for preserving choreography were also solicited. The result of the study led to the design of a robust archival system that can be used for dance preservation by novices as well as experienced people. With, these, it is believed that apart from archival, retrieval will be easy which will lead to a good preservation technique.

Table 2.2. Items that are required or ideal for Preserving Dance Works (Kim, 2012)

Items required for preserving choreography:
Dance notations and scores
Any materials relevant to the choreographer and/or choreography
Any materials self-generated by the dance organizations
Any materials related to the use and preservation of the choreography
Other ideal components and conditions:
Recordings and notation from the inception of choreographic work
Properly housed content
Choreographer's teaching curriculum materials
Enabling and controlling access to the choreography
Dancers' memories of a choreography

The organization Dance America worked at dance preservation under the following mission: Engagement through meaningful programs, convening, and educational opportunities; Advocacy for increased visibility and engagement; research in the dance field and preservation of America's dance legacy. This they did until they merged with DHC. DHC's mission is to document, preserve and create access to the dance legacy of the United States.

Indian dance preservation

Some efforts have also been made to preserve the Indian Classical Dances were found in literature. Cushman, Ellen, Ghosh, and Shreelina (2012) tried to preserve the Indian Classical dances by storing the dance on CDs. This method seems primitive. It is also not stable because CDs can get lost or depreciate very quickly.

Another method of dance preservation used for the Indian dances is the organization of ceremonies. Dance has been essential in furnishing American Indians with a technique for cultural preservation, a religious association, and to function as a community. Out of the conventional ceremonies of the past have developed the different styles of powwow dancing; other customary services have been restored and are presently being preserved. This action has led to younger ones asking questions and people wanting to know about the traditions. Indian powwows keep on picking up prevalence all the nation through these gatherings. Dance, in its numerous structures, keeps on giving a social scaffold to American Indians. Social conservation of American Indian way of life in Denver, Colorado, got a positive lift with an organization made between a nearby school area and a neighborhood Indian training gathering; the aftereffect of this association was a program intended to coordinate an exact depiction of American Indian history with the school educational plan through guidance and reading. This endeavor tells how event organization has served as a positive tool for cultural preservation.

Mallik, Chaudhury, and Ghosh (2011) presented another preservation effort for intangible heritage preservation. It is believed that these cultural aspects are mostly preserved by artifacts. Thus, the use of Multimedia Web Ontology (MOWL) which supports probabilistic reasoning using media properties was proposed for encoding the domain knowledge. This framework includes constructing the ontology using labelled set of training data and using the ontology to automatically annotate new digital heritage artefacts. The annotations are then used as a semantic navigation environment in the cultural heritage repository. The key contribution of this work based on the Indian Classical Dance preservation showed that computing methods can be used to preserve lining heritages. This framework provided a conceptual linkage between multimedia data at feature-level and heritage resources at the knowledge-level. One of the key ingredients in the architecture is a cultural heritage ontology. Mallik and Chaudhury (2009) encoded in a new multimedia ontology representation. Figure 2.4 is a representation of the ontology system developed.

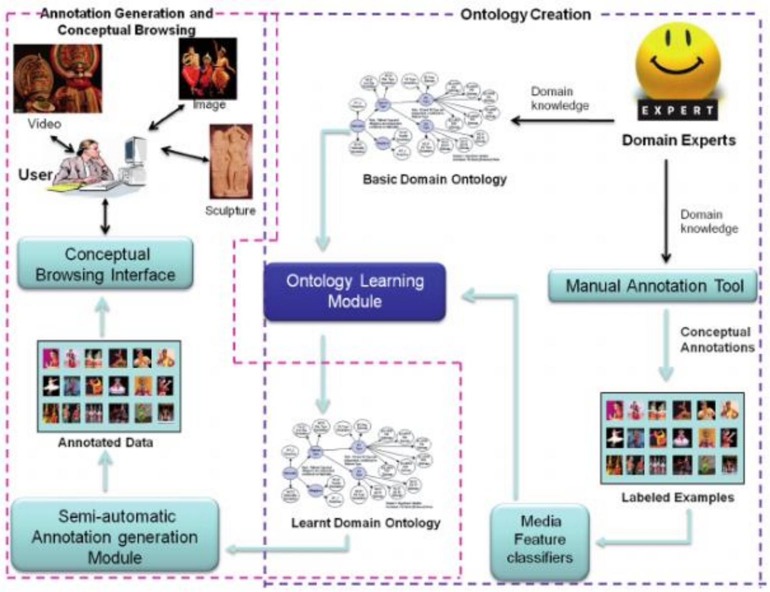


Figure 2.5: Ontology of Indian dance
(Mallik and Chaudhury 2009)

A few instances of the various types of events from ICD space are the accompanying.

1. Spatial events. These are hand gestures, body postures, facial features, and facial expressions.
2. Temporal events. A choreographic succession identified with a move; steps like making a circle, strolling; a grouping of hand motions and body acts that express the expressions of a tune. Classical dance usually has a set pattern of steps following each other. These steps are usually given a name which can be used to reference it. For example, Bhumi Pranam (bowing to the earth), is a choreographic sequence in an Indian Classical dance where the dancer squats, with her knees bent, and uses both palms to touch her forehead after touching the floor.
3. Spatio-temporal events. These helps recognize the different roles played by different objects in a video shot. For example, a group dance usually has a lead dancer.

Artese and Gagliardi (2012) described a project which utilized the definition, implementation, population and search of a register of the intangible cultural legacy of trans-border Italo-Suisse heritage, with the aim to design such a register of such nature for the new heritage paradigm which was proposed by UNESCO. The E.CH.I. Project presented the ICH cataloging card for the inventory of intangible cultural heritage on the web, as a result of its integration in the AESS database that stores information about the oral history of the Italian Lombardy territory. All of these were efforts made at preservation.

The bottom line of Indian Classical Dance preservation is the provision of a collection of dance artifacts and resources, design of an ontology/taxonomy that makes retrieval easy and show casing these cultures in different community gatherings to revive knowledge and interest in the culture.

Chinese dance preservation

Classical Chinese dance is one of the most enticing of visual spectacles. It comprises of movements of exactness, beauty, and control, with some of the most difficult moves—tumbling, jumps, flips, and leaps (Wu, 2015). The moves in this classical dance reflects the martial art techniques from ancient battlefields. These by themselves make old style Chinese move one of the most outwardly great types of move, yet they would be fragmented and shallow without the most significant trademark—an exceptional element called bearing (Hum, 2005). This bearing is at the very heart of classical Chinese dance and it is unique to the culture—a reflection of Chinese civilization. The, Shen Yun Performing Arts, a New York-based group generally is working hard at preserving this tradition by performing classical Chinese dance to different people. This has promoted interest in the dance and has a global phenomenon, selling out some of the world's greatest theatres like the Lincoln Center and the Kennedy Center in the U.S., and the U.K.'s London Coliseum since the program was started in 2006.

A similar preservation style is also used in the preservation of Miao songs. The villagers are often organized to participate in Miao traditional songs and dance. With budgetary help from the nearby government and the Landscape Authority, Miao customary tunes and move classes are held by the town advisory group in the network. Some gifted artists in the town are chosen to contemplate Miao tunes. The original flavor of Miao melodies and moves frequently pull in guests to take an interest and appreciate with neighborhood inhabitants. This is a way that a small community is trying to preserve their cultural heritage. These heritage preservation style follows the traditional method and it is only different because it is aggressive, and the aim is strictly followed. These preservation methods did not really involve the use of technology but arouses the interest in learning the dance in younger generations.

UNESCO in its pursuit of cultural preservation have been able to produce documentaries of some Traditional African dance forms. There is sparse literature about African dance preservation efforts. However, because most African countries are close knit communities, the traditional dances are still performed on occasions which propagates the knowledge of these Traditional African dance forms and motivates the younger generation to learn it. As globalization is affecting local communities, it is becoming tougher to get the younger generations to learn and know about these cultural expressions.

African Dance Preservation

An important step towards African dance was that taken by (Green, 2018). She developed a documentation process called Greenotation. This is meant to provide a means of documenting African dance in a way similar to Labanotation and Benesh Movement Notation (BMN). Figure 2.5 is the Greennotation of Adawe a clap dance for girls in Ghana. She produced a similar greennotation for Agbadza, a war dance from Ghana as well.

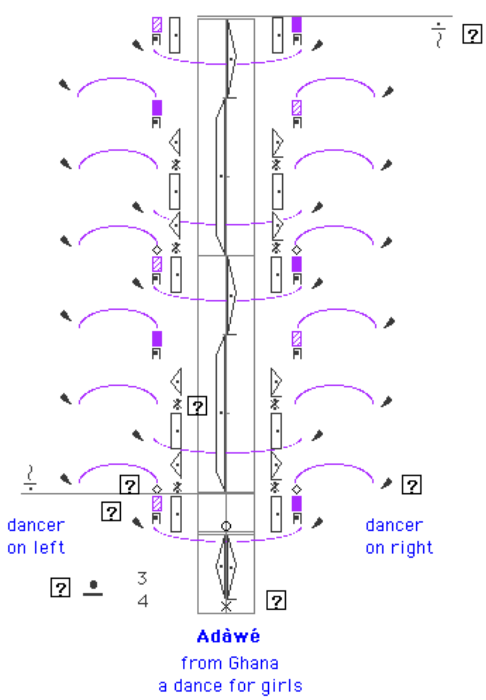


Figure 2.6: Greennotation of Adawe dance (Green, 2018)

Asides from this effort, most preservation efforts for African dances has been in performances in different parts of the world by professional dancers. There are different professional Traditional African dances in Europe, America, Canada and different parts of the work who perform at shows. However, some of this dance performances are limited by available musical instruments, population, language and occasions.

2.6 Technologies used for Cultural Preservation

The application of computer applications and technology in general focusses mainly on improving digitization and documentation of artifacts and sites as well as exploration and reconstruction of monuments (Pavlidis, Koutsoudis, Arnaoutoglou, Tsioukas, & Chamzas, 2007). The use of technology in cultural heritage preservation has become popular with the application of technology to every field. The field of tangible heritage preservation was quick to experience technology integration. The GIS technology became a usual tool for experts involved in cultural heritage preservation such as heritage managers, archaeologists, conservators, and architects (Petrescu, 2007). GIS has been used to build robust information systems for several cultural heritages. These has been used in United Kingdom (Frogat, 2006), South Africa, Australia (Gleby, 2007) Morocco (Kölb, Boussahl, Hostettler, 2003), Japan (Takase, Yano, Nakaya, Isoda, Kawasumi, Matsuoka, Tanaka, Kawahara, Inoue, Tsukamoto, Kirimura, Kawahara, Sho, Shimiya, Sone, & Shiroki, 2006), Germany (Reitz, Haist, & Wigg-Wolf, 2006) and many other countries.

The use of Internet of Things as a tool in Cultural heritage preservation is also growing. The application of IoT architecture to indoor cultural environments like museums or art exhibitions smart are being designed. An example is in its application within DATABENC , the high technology district for Cultural Heritage management founded in Regione Campania, Italy . The integration of smartness is targeted at adding life and motion to what is inherently redundant.

The application of 3D modelling techniques is used often for the reconstruction of physical and natural heritage sites. It is a common place thing to see a miniature models of heritage buildings like museums, historical buildings, and sites that is held in high esteem as part of the story of a community. The development of 3D modelling paved way for the application of Virtual reality in cultural preservation. Not only are models of these heritages constructed, Virtual reality paved way for the construction of Virtual heritages. Virtual Heritage is a cultural heritage within the domain of technology (Rahaman and Tan; 2010) . There are three main domains considered here 3D documentation, 3D representation and 3D dissemination.

3D documentation entails the gathering of accurate information about the site that is being studied. This includes site investigation, taking of measurements, sketching, epigraphy etc. Techniques employed in data gathering includes 3D laser scanning, photogrammetry, stereo-photometry, laser triangulation etc. (Cignoni & Roberto, 2008). The second stage is either the visualization of the object or replication of the object (Rafi, Salleh, Paul, Noraisah, Jun, Hanif, & Mahadzir, 2010). The last domain, 3D dissemination involves the display of the 3D models to users. The effect of dissemination is in making the users experience the heritage object. Examples of these projects include The Great Buddha (Ikeuchi, Oishi, Takamatsu, Sagawa, Nakazawa, Kurazume, Nishino, Kamakura, & Okamoto, 2007), 19th Century of Italian Theatre (Valtolina, Franzoni, Mazzoleni, & Bertino, 2005), Ancient City of Hue, Vietnam (Pugnaloni, Issini, & Minh, 2008), and Anyang Xinyu project (Xinyu, Baoqing, & Chuangming, 2007).

Voinea, Girbacia, Postelnicu, and Marto (2019) developed an augmented reality (AR) application which made it possible to visualize and explore a 3D model of a fortified church. The Project named, Tango, improved did not just apply 3D technology to this heritage but also applied AR technology.

Cultural Computing (CC), which is the application of computer technologies in the field of culture, arts and social sciences is enhancing, extending and transforming the application of computing in cultural preservation (Wang, 2009). The virtual exploration of underwater archeological sites using the VR and AR technologies had tremendous success as a preservation of the site and provision of underwater experience to the users (Haydar, Roussel, Maïdi, Otmane, & Mallem, 2011). Technology and internet boom have made cultural heritage preservation a very booming field to apply technology (Berndt & Carlos, 2000). The availability of many technologies is a match for the various forms of cultural heritages that are available. The field of intangible cultural heritage is also experiencing the use of technology.

The performance of a simulation of Royal Dance of Chu at the palace of Zhongshan was used to capture ancient crowd activity for the preservation of that heritage using VR (Cheng, Peng, & Sun, 2006). Movements were collected using motion captures and these were integrated in the reconstruction to imitate the cultural space and the activity that transpired there. Film documentaries are commonly used for the preservation of intangible cultures. This was the traditional way of preservation intangible culture. Documentary films are used to capture culture and important moments in history. Festivals, rituals, songs, dances, religious rites etc. are being preserved for the future generations using films. There are several documentaries available about events in history and the way of life of people.

With the advancement of technology came the web documentaries. While they are also some sort of film documentaries, they have the element of interactivity embedded in them. Highrise: Out my window (http://outmywindow.nfb.ca/#/outmywindow), Collapsus (http://www.collapsus.com/), Arte.tv's (https://www.arte.tv/fr/), Prison Valley (http://prisonvalley.arte.tv/?lang=en) are examples of web documentaries. Highrise: Out my window is a documentary about people leaving in the high-rise buildings around the world. It featured 49 stories from 13 different cities which are told in thirteen languages. It discusses the culture, their backgrounds, challenges, what this place means to them and essentially a documentation of their experience. While web documentaries are like film documentaries, they add the beauty of individuality and experience. In a film documentary, the observer is learning from an experience and watches the story unfold. In a web documentary, the person interacting with it moves from being an observer to a participant. Users can determine the path they want to follow because the experiences are divided into different modules. The observer is experiencing it. Web documentaries have a lot of other technology like VR, flash videos, and 360 degrees views embedded in them.

Another method that can be used to digitize intangible cultures is the use of website. While several of these methods can be used together to achieve the same aim, individual digitizers usually choose the most appropriate methods for their content. Websites can be used to keep videos, animations, photographs, text, etc. that will provide contents to people who wish to learn about the culture or experience the culture. A website will remain available as long as it is hosted. Websites are being used a lot as a basic way of broadcasting all the other methods of digitization and preservation method. Examples of cultural dance preservation sites are:

- https://danceinteractive.jacobspillow.org/themes-essays/african-diaspora/,
- http://new.danceheritage.org/html/treasures.html,
- https://www.thirteen.org/freetodance/timeline/index.html,
- https://blackballerinadocumentary.org/.

Photography is the easiest and most used for intangible cultural preservation. Whether the photos are taken by a professional or not does no matter as much as the fact that the content is made available for people later. With the development of digital cameras and now high-resolution phone cameras, pictures are easier to use than before. When it was the analog camera, there is need to digitize the photos by getting them scanned and saving them away securely. The challenge with that generation of photos is that the hard copies needs to be preserved under a very strict atmospheric conditions lest they are ruined. The limitation with photos is that they are still images, which means they only provide spatial information and not the temporal information. This limitation is reduced with the development of the live photo technology. The iPhone live photos capture three seconds of movement and sound in the photo. It captures 1.5 seconds before you take the shot and 1.5 seconds after it. This way the photo is no more still but a little dynamic.

Photographs can be printed in books, stored in the cloud or in order forms to be retrieved when needed. The use of photography is very common in individual family lives.

2.7 Deep learning Techniques

2.7.1 Optical Flow

Videos are collection of images moving in a very fat sequence. Since the data that will be used for the study are videos that will be converted to images, it is important that the motion between one frame to another is captured in the action. This will be achieved by using the optical flow. Optical flow can be defined as the pattern of apparent motion of objects in images from one frame to the next in the sequence which is caused either the object or camera movement. It is 2D vector field. Each vector is a displacement vector.

2.7.2 Action Recognition Algorithms

Convolutional Neural Network

The overall agenda of the computer vision domain is to enable machines with the ability to view the world as humans do. Added to this ability is to be enabled to perceive the world the same way humans do and be able to use the knowledge to perform tasks such as image recognition, video recognition, image classification natural language processing, etc. CNN has been a primary algorithm that drives this world. The success of CNN is a reason why deep learning has become popular. LeCun, Bengio, and Hinton (2015) asserted that deep CNN is a breakthrough milestone in image, video, audio, and speech processing. CNN takes and processes images as tensors which are essentially matrices of numbers with additional dimensions. When a computer is given an image, what it sees is an array of numbers, a matrix (Figure 2.6). b

Figure 2.7 Images as seen by the Computer.
Retrieved from https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/

The size of this matrix is determined by the size of the image. A CNN is made up of an input layer, output layer and a multiple of hidden layer. The CNN consists of one or more convolutional layers which are then followed by one or more fully connected layers. A simple architecture of a CNN is in figure 2.7.

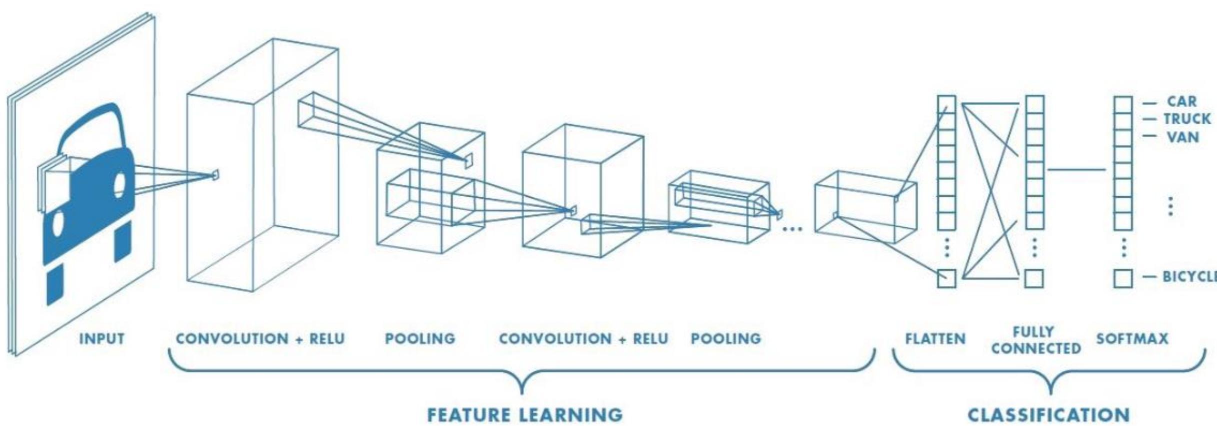


Figure 2.8: A sample CNN.
Retrieved from https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

The input which is an image is read as multiplication of number of images, image width, image height and image depth. The higher the resolution of the image, the bigger this number gets. The convolution layer performs convolutions of these array. Convolution is a mathematical operation that involves two functions producing another function. Here, a matrix in the image is multiplied by another matrix to produce another matrix of lower dimension. In convolution this means that all the pixels in the matrix interact with one another and produce another pixel. The convolution layer convolves the image and pass the result of the convolution to another layer. The result is convolved feature. The pooling layer reduces the dimension of the data by combining the neuron clusters that are outputs of one layer into a single neuron in another layer. Thus a 6 by 6 matrix can be reduced to a 3 by 3 matrix. Although there are different types of pooling the most commonly used are max pooling and average pooling. In max pooling the maximum value in the cluster of neurons is used in to represent the cluster in the next layer. Average pooling on the other hand takes the average of a cluster as the representation of the layer. The fully connected layer performs the task of connecting all the neuron in a layer to the neurons in another layer. This connection is what is used to classify the image.

In the field of action recognition, CNN can be divided into 2D CNN and 3D CNN. 2D CNN are mostly used for action recognition in images while 3D CNN are mostly used for videos, but each type can both be used interchangeably in different scenarios. Karpathy, Toderici, Shetty, Leung, Sukthankar and Li (2012) performed large-scale video classification using a dataset of 1 million YouTube videos consisting of 487 classes. The model performed excellently well on the UCF-101 Action Recognition dataset at 63.3% accuracy compared to UCF-101 baseline model which was 43.9%. Wang, Qiao and Tang (2015), in trying to understand human actions in video developed a novel form of video representation called trajectory-pooled deep-convolution descriptor (TDD). This new feature extraction method was applied to a CNN that was used for action recognition. When applied to the UCF-101 dataset, it achieved a 91.5% accuracy on recognition.

Simonyan and Zisserman (2014b) designed a two-stream architecture for video classification based on a CNN. The first stream is a convolution of the spatial part of the video. The spatial part is a form individual frame sequence about the scene and the object. This is practically an image. The second stream targets the temporal part of the video. The temporal part is a form of motion across the frames. This motion was captured as an optical flow. This two-stream model achieved 88% accuracy on the on the UCF-101 dataset. The result showed that training on optical flow is a good for action recognition in videos than just using images alone. Feichtenhofer, Pinz and Zisserman (2016) improved the two-stream network by not fusing them at the SoftMax layer and doing this in the last convolution layer. The spatiotemporal fusion of the video snippets achieved a 92.5% accuracy on the UCF-101 dataset. Fernando, Anderson, Hutter and Gould (2016) on the other hand used hierarchical rank pooling as a method encode video sequence for activity recognition in a video. When this model was evaluated on state-of-the-art activity benchmarks it performed well. The model achieved 91.4% accuracy on UCF101, 66.9% on HMDB51 and 76.7% on Hollywood2 datasets.

Girdhar, Ramanan, Gupta, Sivic, and Russell (2017) designed introduce a new way of video representation for action classification that combines local convolutional features across the entire spatio-temporal extent of the video. It aggregates a set of action primitives over the appearance and motion streams of a video which is then used for video classification. The ActionVLAD achieved a 93.6% and 69.8% accuracy on the UCF-101 and HMDB51 datasets respectively. Kar, Rai, Sikka, and Sharma (2017) introduced the concept of adaptive pooling for temporary pooling frames. This method learns to pool discriminative and informative frames and discard most of the non-informative frames in a single temporal scan of the video. This method of mean pooling achieved a 93.2 % and 66.9% accuracy on the UCF-101 and HMDB51 datasets. Cherian, Fernando, Harandi, and Gould (2017) instead of discarding temporal order of frames in the pooling layer, introduced a novel pooling method called generalized rank pooling. This pooling method takes as its input the features from the intermediate layers of the CNN that is trained on tiny sub-sequences and then produce as the output the parameters of a subspace that provides low-rank approximation of the features and preserve the temporal order as well. The application of the generalized rank pooling achieved a 93.5%, 72.1% and 73.7% accuracy on the UCF-101, HMDB and JHMDB datasets respectively. The different variations in the application of CNN in action recognition is endless. Classic CNN networks are LeNet -5, AlexNet, VGG-16, ResNet. These classic networks have been published online and can be use by anyone. The advantage here is that one can build on the strengths of these networks to make a new network.

LeNet-5

LeCun, Bottou, Bengio and Haffner (1998) developed the LeNet CNN architecture (Figure 2.8). It is a feed forward Neural Network that is made up of five alternating layers of convolution and pooling layers which are followed by two fully connected layers. It made use of TanH as its activation function. It was developed on the idea that neighboring features are correlated to each other and they are distributed across the entire image. This solved the problem posed by the traditional fully connected Neural Network. It was the first CNN architecture which was able to automatically learn from raw pixels and reduced the number of parameters. This made it very famous and popular. It became the template for constructing CNN models. The setback for this network is that the activation function, tanH is extremely slow.

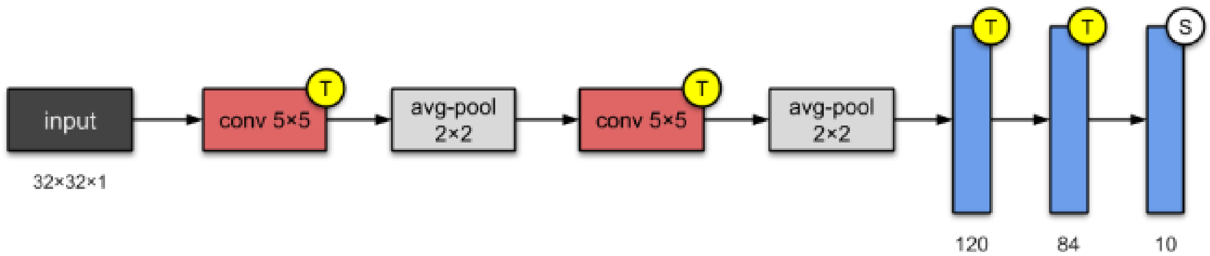


Figure 2.9: LeNet-5
(Raimi, 2019)

AlexNet

AlexNet was the first deep CNN architecture with enhanced learning ability and groundbreaking results. At the time of development there were advancements in hardware features which enabled it to utilize two NVIDIA GTX 580 GPUs that made processing and feature learning faster. It was developed by Krizhevsky, Sustskever, and Hinton (2012). The network was the first to implement Rectified Linear Units (RELUs) as activation function. The neural network has 60 million parameters and was made up of 8 layers – 5 convolutional layers and 3 (figure 2.9).

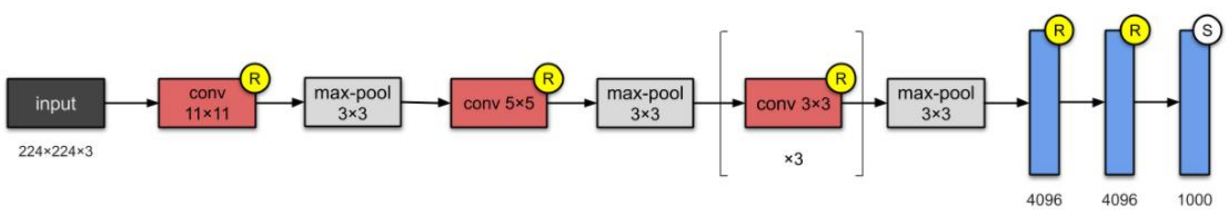


Figure 2.10. AlexNet Architechture
(Raimi,2019)

VGG-16

This network was invented by Simonyan and Zisserman at the Visual Geometry Group (VGG) in 2014 (Simonyan and Zisserman, 2015). This network went deeper that AlexNet by going to having 16 layers, but in this case small size filters were used. Thus, they replaced 11 by 11 and

5 by 5 filters with 3 by 3 filters. The use of small filters reduced the computational complexity by reducing the number of features. The ReLU activation function was used. It is simple, homogenous and have increased depth. However, it is difficult to deploy it on a low resource system and it is computationally expensive. Figure 2.10 presents a visualization of the network.

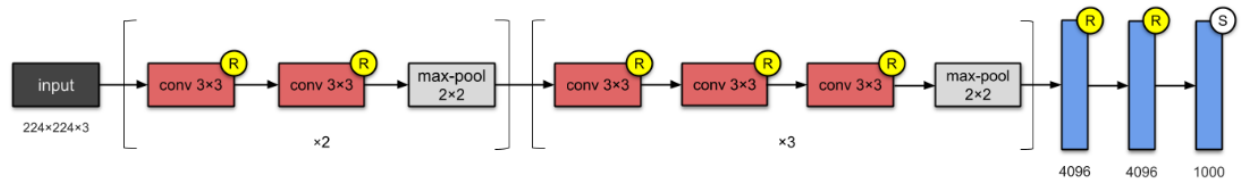


Figure 2.11 VGG-16 Architecture
(Raimi, 2019)

Another network called VGG-19 has been developed by the same authors and it has 19 layers.

ResNet

He, Zhang, Ren, and Sun (2016) developed the ResNet as an improvement over AlexNet and VGG. There are three versions 50 layers, 101 layers and 152 layers deep. It achieved less error on image classification task. The performance of ResNet on image recognition established representational depth as centrally important to visual recognition endeavors. Figure 2.11 is the architecture of ResNet-50.

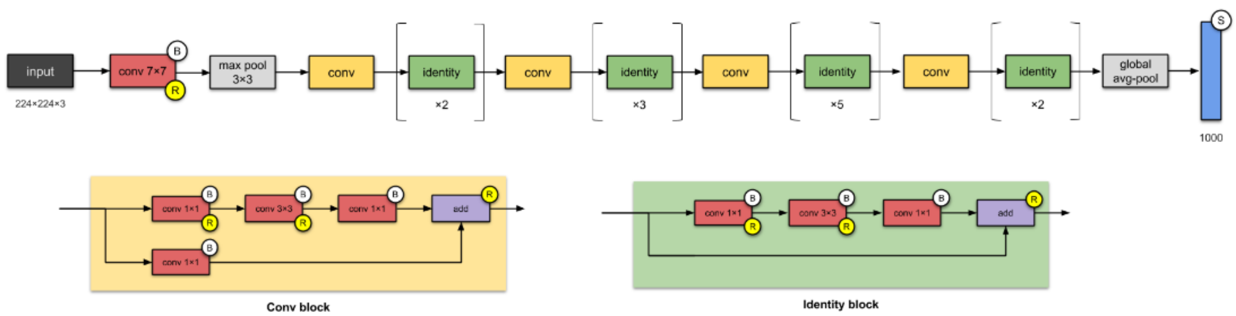


Figure 2.12: ResNet Architecture
(Raimi, 2019)

Inception

There are different versions of the inception frameworks. There are inception V1, V2, V3, V4 and inception-ResNet. The inception-v4 was enabled for 43M parameters while the exception-

ResNet was on 56M parameters. The idea of the inception network is to reduce the cost of computation by deep networks and not affect the generalization. The inception-ResNet combined the power inception block and residual learning. Inception-ResNet converged more quickly that all the other versions. Thus, the strength of this network is that it achieved good results in a lesser time. The networks were designed by Szegedy, Liu, Jia, Sermanet, Reed, Anguelov, Erhan, Vanhoucke and Rabinovich (2015).

Kumar, Kiran Kumar, Sastry, Kiran, Anil Kumar, and Prasad (2018) used a multistage 3D CNN for the classification and identification of classical Indian dance. The model consisted of one input layer, four convolutional layers, two stochastic pooling layers, one dense and one SofMax output layer. Figure 5 provides the architecture of the CNN. Rather than use joints of the dancer to learn dance poses, Kaushik, Mukherjee and Lall, (2018) used a CNN based architecture for designing a novel pose signature for a sequential dance learning framework for the classification of Indian classical dance. The pose estimator was based on the output of the 3D CNN features generated. It is on this estimation that other frameworks were applied.

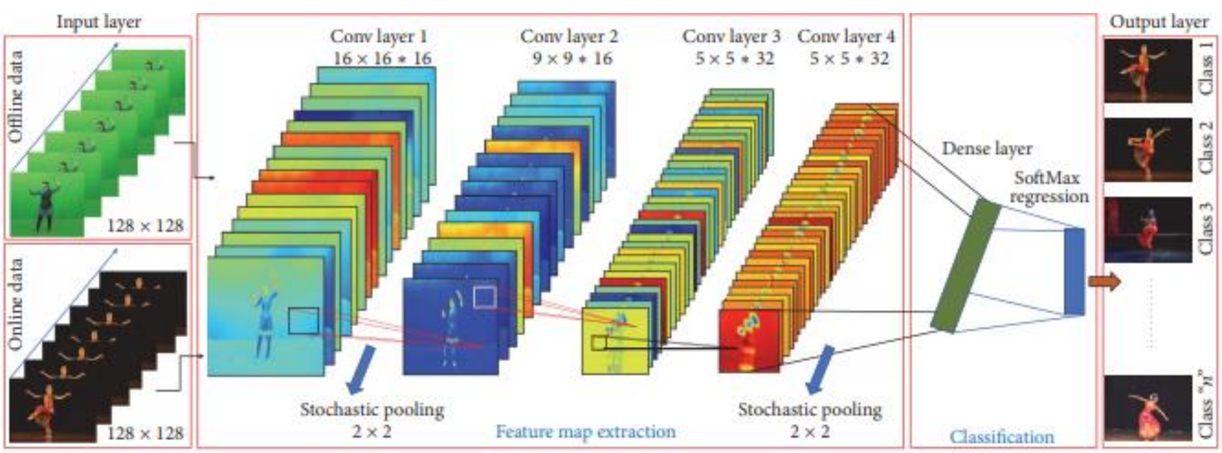


Figure 2.13: 3D CNN Architecture for Classification of Indian Dance.
(Kaushik, Mukherjee and Lall, 2018)

Bakalos, Rallis, Doulamis, Doulamis, Protopapadakis, and Voulodimos (2019) identified and simulated moving behavior in dance choreographies using CNN. The architecture was used to identify the choreographic poses in the dance video sequence. Costa, Soares de Oliveira and Silla (2017) used CNN for music classification using spectrograms. Castro, Hickson, Sangkloy, Mittal, Dai, Hays, and Essa (2018) used a CNN to address the regression of human joint location estimation. The CNN was used to extract the 14 essential body points in the dance video. The extracted images were then used for classification.

Recurrent Neural Network

Recurrent Neural Networks (RNNs) is an interesting twist to basic neural networks, this is because they are designed to take a series of input with no predetermined limit on size, ie it can take in a larger input in terms of features and inputs. This neural network is enabled to remember the past when deciding. So, it is not just about the current image but can 'remember' the previous images too. In this network type there are connections from the previous node to the next along a temporal sequence. This attribute makes it efficient for video data. Rather than try to use deep CNN networks to learn the temporal part of a dataset for deeper accuracy, the RNN is already enabled to capture temporal sequence. Figure 14 is the architecture of a traditional RNN. The RNN uses previous outputs as input for the next layer.

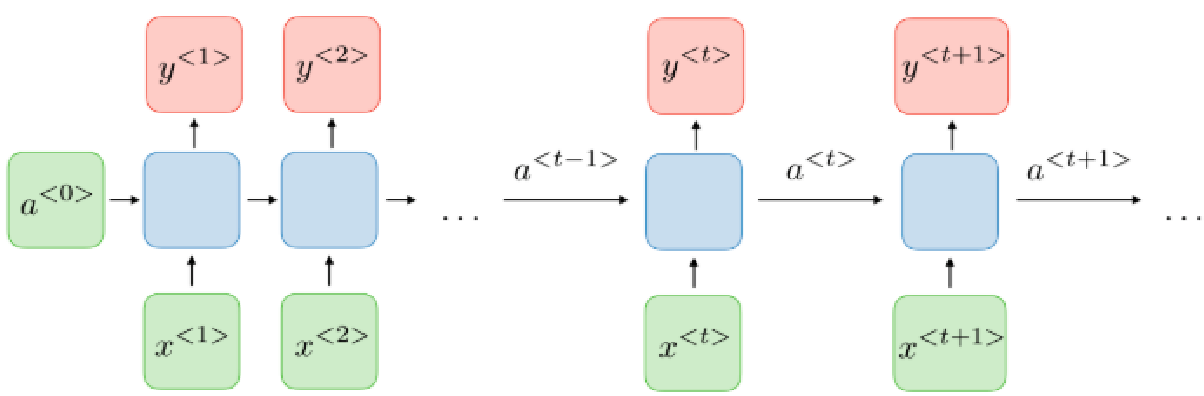


Figure 2.14: RNN Structure.
Retrieved from https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks

The beauty of this neural network is its application to datasets with time series. It is able to isolate important patterns in the sequence without having overlaps. The drawbacks are that its computation is very slow, and when considering features that are very wide apart in the sequence; it cannot perform better than a CNN because it is unable to make the connections. RNNs are usually applied in music generation, sentiment classification and name entity recognition. Sigmoid, Tanh and ReLU are the most commonly used activation functions. Long Short-Term Memory

(LSTM) and Gated Recurrent Unit are the commonly used types. Both models solved the problem of vanishing gradient that is experienced when traditional RNN is used. LSTM has a ‘long term’ memory which enabled it to be capable of processing complex sequential information (Zhuang, Qi, Duc Kieu & Hua, 2019).

Zhao, Ali, and van der Smagt (2017), carried out action recognition by combining RNN and CNN. This model achieved a 14% raise in accuracy compared to state-of-the-art methods. Pienaar and Malekian (2019). achieved a 94% accuracy in human activity recognition by combining LSTM and RNN deep neural network architecture. Ma, Chen, Kira and AlRegib (2019) applied Ts-LSTM for exploiting spatiotemporal dynamics for the essence of action recognition. Krishnan, Prabhu and Babu, (2016) made use of LSTM in action recognition. The endeavor achieved an accuracy of 87.2% when applied on the Penn action dataset. The popular thing is combining different frameworks to g l the information that is needed for a great output.

2.7.3 Action Generation and Algorithms

Generative Adversarial Networks

GANs are an emerging technique for generating actions from given that takes on the form of what the machine was fed with. It is used for both semi-supervised and unsupervised learning. This was proposed and achieved by Makhzani, Shlens, Jaitly, and Goodfellow in 2014 but did not become popular until 2016. It has been used for various augmentation, transfers and deep fakes. The framework for GANs is that it is built on two different neural networks. The first is referred to as the generator while the second is referred to as the discriminator (figure 2.14).

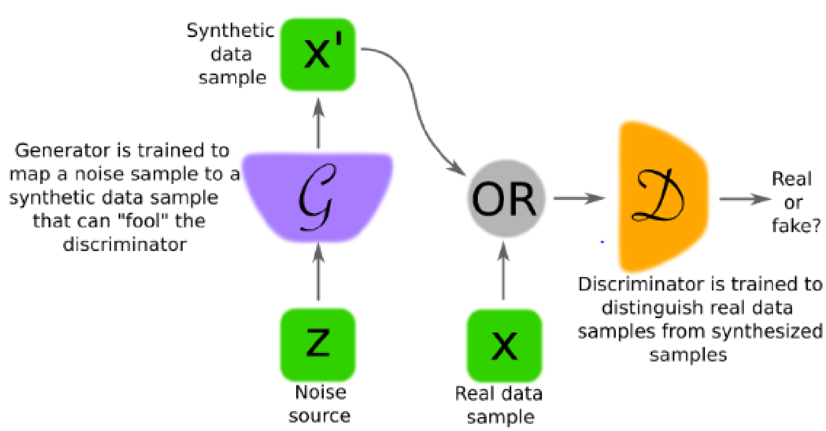


Figure 2.15: GANs framework
Source: Creswell *et al.*,., 2017.

The generator tries to generate images that are fake but like the data given. The discriminator on the other hand is playing a detective role by telling the real images apart from the fake images generated by the generator. This game of between the generator (thief) and the discriminator (detective) continues for a while until the generator can generate an image that the discriminator is unable to identify as fake.

Through this method the machine can produce a replica and, in some cases, enhanced images from a given data which is hard for human to tell apart whether they are real or not. GANs has being used severally in developing photorealistic images, transfer features from one image to another, generate deep fake images and videos, and even in music generation. This study will use the same artificial intelligence to generate Traditional African dances.

2.7.4 Pose Estimation Algorithms

Pose estimation is a part of computer vision that estimates the configuration of the human body parts from images, videos or data captured by sensors. This information has applications in Augmented reality, motion analysis, security, healthcare, virtual reality, gaming etc (Moeslund & Granum, 2001; Moeslund, Hilton, & Kruger, 2006). There are several advances in this area of study in recent years. There are different models used in pose estimation but the most common is the skeleton-based model (Ji & Liu, 2009). It has a tree structure with nodes at each body part. The limitation of the skeleton-based model is its limitation in representing shape and texture information (Poppe,2007). Challenges in pose estimation includes inability to represent occluded body parts, insufficient training data and depth issues. Two major algorithms have been very successful in the recent time.

Deep Pose Algorithm

Toshev and Szegedy (2014) in their work applied regression to the problem of pose estimation. A function $\psi(x;\theta)\in R2k$ was trained and use predict the absolute image coordinate as represented in the equation below. $y*$ is the pose prediction.

$$y*=N-1(\psi(N(x);\theta))$$

The algorithm architecture (Figure 2.15) shows that the first layer takes in the image at a predetermined size. A deep neural network with seven (7) convolution layers was used perform the regression analysis which outputs the joint localization. Another seven (7) layers of regression was applied to it and this outputs the pose estimation. This algorithm was based on Krizhevsky, Sutskever, and Hinton (2012) algorithm.

Figure 2.16: Deep pose Algorithm Architecture. Source: Toshev and Szegedy (2014)

Open Pose Algorithm

Cao, Hidalgo, Simon, Wei and Sheikh (2021) recently progressed on their Open Pose algorithm which was designed on the CMU Panoptic Studio dataset. The pipeline of the method used in open pose is presented in Figure 2.16. The system takes an input and produces a two dimentional location of all the key body joints. The second stage is to predict the confidence of the body part location and then use that to calculate association between the parts. The confidence map derived at this point is then used to produce the pose estimation. This algorithm works well on body parts that are not joints such as eyes and mouth.

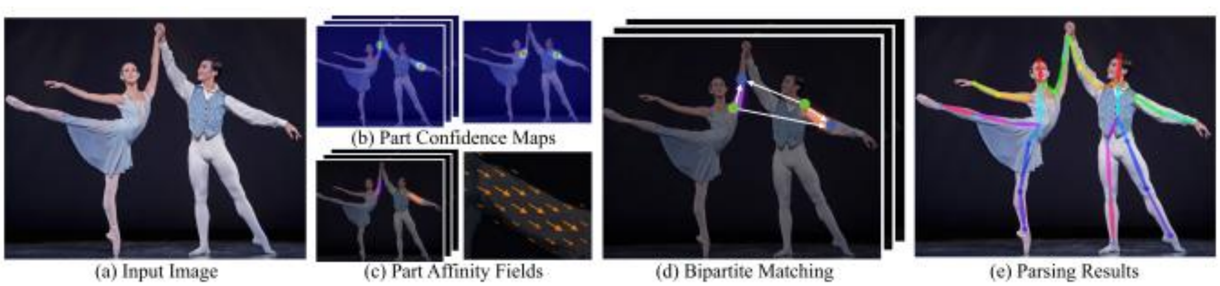


Figure 2.17: Overall Pipeline of Open Pose Algorithmn. Source: Cao, Hidalgo, Simon, Wei and Sheikh (2021)

2.8 Adowa

2.8.1 History

The Adowa dance is a dance of the Akan ethnic group of the Ashanti region in Ghana. It is one of the numerous traditional dances in Ghana. The Adowa dance is named after *Adowa* (Antelope). The dance evolved from the movement of an Antelope. History has it that at a time in the Ashanti region, the queen mother Abrewa Tutuwa was sick and the gods require that an antelope should be sacrificed so she can recover (Ampomah, 2014). After a long search an antelope was found and kept in the cage before its final day of sacrifice. But right in that cage, it was noticed that the antelope was making movements that can termed as dance. People began to display the different moves that they have learnt from the animal. This movement eventually evolved to be called Adowa dance. Although Adowa was traditionally a funeral song, it has evolved into dance that is performed at all forms of social events now (Green, 2012). The gestures used in the dance varies with the event.

2.8.2 Costume

Adowa dancers are usually dressed elegantly because of the dance's history. The dancer is 'a queen' and so he or she is expected to dress up to show that fort. Sometimes Adowa dancers are dressed in black, this is usually when the dance is performed at funeral services. The dancers are usually dressed in Kente or adinkra material. For the female the cloth is usually just tied round the body. Male dancers wrap the clock around their torso and leave the left shoulder bare. Dancers wear gold and bead ornaments on their hair, arm, wrist, neck, ankle and sometimes leg. Dancers also have scarves just hanging from their waste or in their hands. This scarf is also used in performance. Elegance is the theme here. Dancers may and may not wear footwears. Figure 2.17 shows Adowa dancers in their costume.

Figure 2.18: Adowa Dance Costume.

2.8.3 Poses

The Adowa dance is characterized by hand and feet gestures. The feet are usually sliding from one side to the other. The hand gestures on the other hand has different meanings. Figure 2.18 presents some dance poses.

s

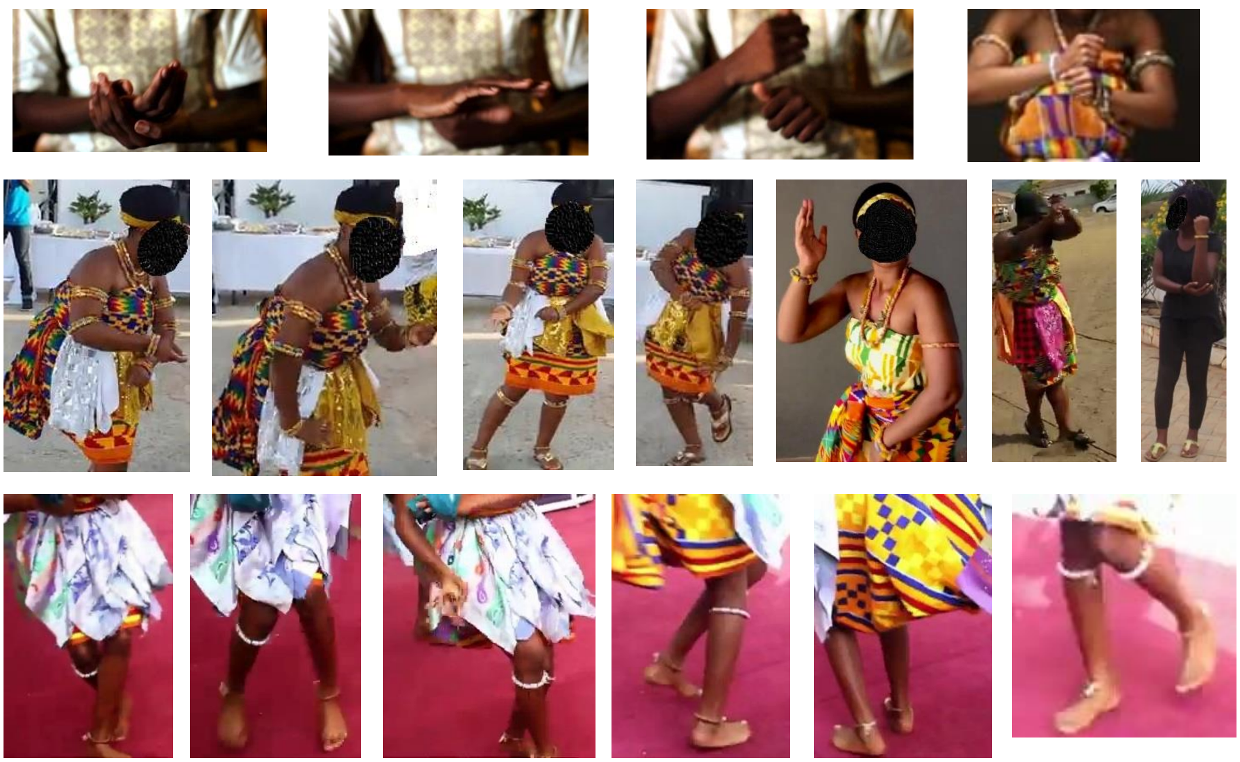


Figure 2.19: Adowa Dance Poses

There are different hand gestures in this dance and they all have individual meanings. Examples of such hand gestures according to Tuffour (2020), and Agyekum (2010) are presented in Table 3. The use of the hand gestures is very important for communcation in the Adowa dance Ghana

News-SpyGhana.com. (2012, May 4). Dancers are not just communicating with the drummers but with the audience watching the dance as well.

Table 2.3: Some Adowa hand gestures and their meanings.

Hand gesture	Meaning
Fist pounding that is consistent with thedrum beats followed by a finger slap.	The dead person is one and loved.
Fist clenching over stomach with crossed arm.	'I am an orphan'
Hand clasped behind back.	'I am lonely'
Left hand flex to the left, followed by right hand to the right and bringing both arms to the chest crossed.	'The whole land is under my control'
Using the thumb and middle finger to form a closed ring.	'Are you married?'
Index finger in the mouth.	Telling the female dance partner 'You are beautiful'
One hand into an upward facing palm of the second palm in a consistent up and down movement.	'Please forgive me.'
Fist held up with second palm wrapping around elbow.	Salutation
Both hands on head.	'I am in trouble'
Throwing up the arms and clasping the hands.	Expression of joy
Placing two palms against one another vertically.	There is peace!
Dancer falls into another person's arms or lap.	'I depend on you'

2.8.4 Drumming

The Adowa dance like most of the African drums have specific drums used for them. In the case of Adowa, five (5) drums are used in the ensemble. These are Atumpan (Figure 2.20a), apentemma (Figure 2.20b), donno (Figure 2.20c), petia (Figure 2.19d) and brenko (Figure 2.20e) (Arthur, 2006; Anku, 2009).

The Atumpan drum is the major drum, it is called the talking drum of the Akan people. The drum is carved out of a tree and placed on stands. It is beaten with two v-shaped wooden sticks with one side of the v longer than the second. The higher pitched drum stands on the drummer's right. Apentemma is a sonorous drum. It is played with the hand. Donno is an hourglass shaped double-headed drum. A side of the drum plays simple duple rhythms and the other plays cross rhythms. The petia drum is a tall and narrow tenor drum. This shape enables it to produce rounded sound. It is played with sticks. Brenko is an open single-headed goblet drum. It is played the fingertips and palm alternately. Two metal bells called Dawuro (Figure 2.20f) are used to accompany the beats. A gourd rattle known as Ntrowa (Figure 2.20g) is usually used. There three major techniques used in the Adowa drumming, these are:

a. Straight stick technique: this is usually used with Apentenma, petia and donno. The donno drum requires more techniques from drummers because it involves squeezing and releasing the throngs alternatively.
b. Hook-shaped stick technique: For this technique the V-shaped or curved sticks are used. This technique is used with Atumpan and Donno.
c. Hand technique: here the drum is played with hands.

(a) Atumpan (b) Apentemma (c) Donno

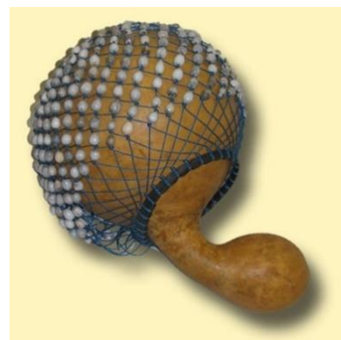
Figure 2.20: The Adowa drum ensemble

Figure 2.20 continued

(d) Petia

(e) Brenko

(f) Dawuro



(g) Ntrowa

2.9 Bata

2.9.1 History

The history of the Bata dance is more of mythology than factual history (Osanyin, 1996). One thing that is clear is that the dance is attributed to Sango the God of thunder and lightning. Sango was also a former king of the Oyo Empire. Dancers imitate Sango's performance. This dance is energetic, dancers use all their body parts to communicate. The rhythms of the Bata drums are performed to by the dancers even though it sounds noisy and confusing to people who have no knowledge of the drumbeats. The lead drummer usually plays texts that usually challenges the dancers. Dancers also challenges themselves, sometimes dancing in pair. This artistic performance usually brings out the beauty of the drumming. The language of the Bata drum can only be decoded by the expert.

2.9.2 Costume

The Yoruba dance has no special costume, dancers are usually dressed in Yoruba traditional wears. However, dancers usually wear low footwears (modern) or none; this is because of the high movement and energy associated with the dance. There is even a common saying that 'A kii fi bata jo Bata, eni to ba fi bata jo Bata, a jo batabata ni.' This means we do not dance Bata wearing a shoe, and that anyone that dances Bata wearing a shoe will not dance it well.

2.9.3 Poses

Dance performance is usually initiated with a call from the Bata drum. The dance is an interpretation of the Bata drum speech. There are several sequence of dance poses. It is characterized by a lot of energetic strikes, swayings, and fast movements. There is a sequence called abula – gbakan. This sequence is a call and response situation. The drummer makes a call called 'abula' and the dancer responds with 'gbakan'. If the drummer calls for two abula, the dancer will respond with two gbakan. Abula is a form of food in Oyo region, the beats are mainly the drummer requesting for serving. Another dance sequence is called 'Furajiga', this sequence is an explanation of how a warrior escapes hidden traps on the battlefield. The 'furajiga' sequence is as shown in figure 2:21.

Figure 2.21: Furajiga dance sequence.

Figure 2.21 continued

2.9.4 Drumming

A Bata is shaped like an hourglass with one side that is larger than the other. Batas are used were mostly used in religious or semi-religious settings by the Yoruba tribe in southern West Africa, as well as by worshippers of Santeria in Cuba. The Bata drum ensemble is made up of three drums as shown in figure 2.22. The biggest of the three is referred to as Iyá-ilu (the mother drum – literal translation). This drum leads the ensemble. It plays complex and long patterns and communicates with the other drums. The second drum Itótele, is smaller than iyá ilu and they have similar tones. The tone of these two drums are made dull by a wax-like material called ‘Ida’ that was smeared on the larger heads of the drum called enu (mouth). The third drum can either be one called Okónkolo or a collection of three small drums tied together.

The collection of three drums referred to as Omele (individually they can be referred to as Omele ako, abo and kudi). Some few people make these a collection of just two drums. The Omele, is high pitched and speak more fluently than the other drums in the ensemble. All three drums are carved from solid wood called Omo, they are not built on staves. The drums dictate the performance of the dancer.

The drums are played professional drummers performed at religious celebrations that are associated with four related deities namely Sango, Oya (Sango's first wife), Egungun (Sango's younger brother) and Sapana (Sango's elder brother). According to Yekini-Ajenifuja (2014), there are five techniques used in Bata drumming. These techniques are:

a. Stick and hand -the big side of the drum is played with hand while the small side is played with stick.
b. Stick and stick- This technique is only applicable to Omele. Two sticks are used to play the faces.
c. Hand and hand – this technique is applicable to all the drums. In this case. The drums can either be played vertically or horizontally. Omele is the only one that must be played vertically.
d. Muting technique: In this technique, a black substance on the side of the drums are either used to change the tone quality or amplify the sound of the drum.
e. Bell technique: For this technique, small bells are tied to the edge of Iya-ilu. The bells are primarily used to resonate and color the music produced.

Figure 2.22: The Bata Ensemble

2.10 Swange

2.10.1 History

Swange was believed to have evolved from Ange by Utuku Agire in 1897 (Tsevende, Agber, Iorngurum & Ugbagir,2013). Swange is a spirited nonritual dance of the Tiv people of the middle belt zone of Nigeria. It is popular performed at all Tiv events (Agber & Ingyoroko, 2012). The dance despite its seemingly heronous outlook has survived religious and western influence (Geri, 2012).

2.10.2 Costume

Dancers are typically dressed in Tiv national attire which is a white and black stripped cloth like that of the zebra. Dancers also wear beads around their arms, hair, ankle, wrists etc.

2.10.3 Poses

The Tiv Swange dance is mainly characterized by fluidity in body movements and swaying of the waist, dancers mimic the flow of river Benue (Manta, 2006). Figure 2.22 displays some Swange dance poses.

2.10.4 Drumming

Swange music is accompanied by three main instruments; Gbande (drums), Kwen (Gong) and Gida (trumpet). Gbande is carved from mostly Gbayee wood and covered with animal skin. It can be single or double sided and are played with hands. Several Gbandes can be played simultaneously. Gida is a wind instrument made from the combination of wood, brass, mild steel and fabrics. Amazing sweet sounds are produced by lip vibration on the cup of the instrument. Kwen is a metallic musical instrument. This gong can be of different sizes. Sound variations are made by tapping on different parts of the gong. A typical Swange band consists of a low-pitched bass Gbande, two smaller high-pitched Gbande, a Gwen and Gida. Figure2.23 displays the instruments.

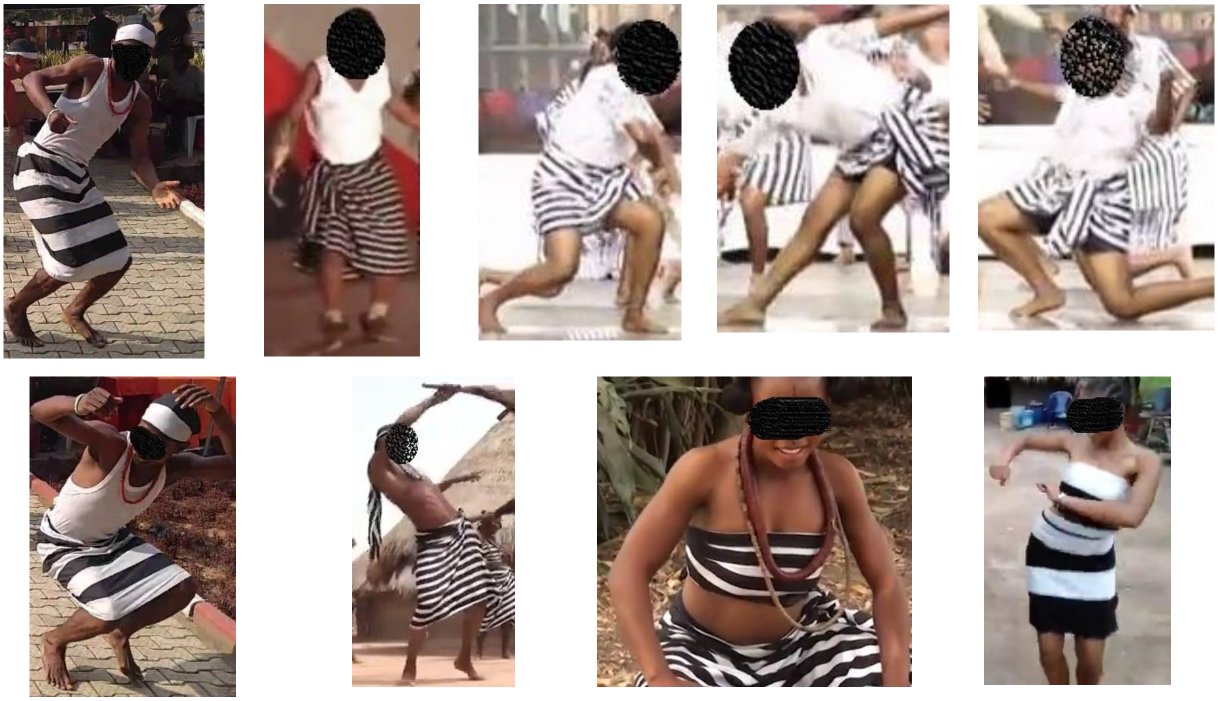


Figure 2.23: Swange Dance Poses

Figure 2.24: Swange Musical Instruments

2.11 Sinte

2.11.1 History

Sinte dance is a dance with its origin from the Boke and Bofa regions in the northwest region of Guinea. It is mostly referred to as the West African Dance. The song is a ritual song of initiation performed by adults to inform the people of some trials they may encounter as they grow. The dance is very popular because of its melody and diversity in movements. Sinte like many other traditional dances are no more limited to performance at initiation rites, it is now performed at many occasions for entertainment purposes.

2.11.2 Costume

Sinte dance does not have a special costume. Dancers usually wear African prints and sometimes are dressed like hunters.

2.11.3 Poses

Sinte dance poses involve swaying from side to side and lifting of the legs at the knee. The dance is a sort of rhythmic chaotic situation.

Figure 2.25: Sinte Dancers

2.11.4 Drumming

The traditional drum used for Sinte dance is a hardwood drum called Krin, however it has been adapted to djembe, a very popular drum in west Africa.

Figure 2.26: Sinte Drums.

CHAPTER 3. METHODOLOGY

The details the method that was used for this study will be discussed in this chapter. Information about how data was collected, the sample, population, study design and how the study models were evaluated is discussed.

3.1 Research questions and hypothesis.

The study provided answers to the following research questions.

a. Can YouTube videos be used to generate sufficient data for Traditional dance classification and modelling using deep learning techniques?
b. Are deep learning techniques suitable for intangible cultural preservation?
c. What deep learning technique(s) will identify and classify Traditional African dance poses?
d. Can deep learning techniques generate a ‘tangible’ dance model of Traditional African dance from videos?

The hypothesis that the study investigated are:

a. YouTube videos will be enough to generate hundred thousand images for dance classification study.
b. Convolutional Neural Network (CNN) will classify Traditional African dance at 95% accuracy.
c. Generative Adversarial Networks (GANs) will generate a Traditional African dance from videos at 95% accuracy.

3.2 Research design

3.2.1 Dance classification design

The dance classification was carried out following a typical video and image classification design as shown in Figure 3.1. The dance videos were extracted into images using OpenCV. Features of the images were extracted. After this, different models of Convolutional Neural Network (CNN) algorithm were then applied to the dataset. The classification models were then evaluated using different metrics.

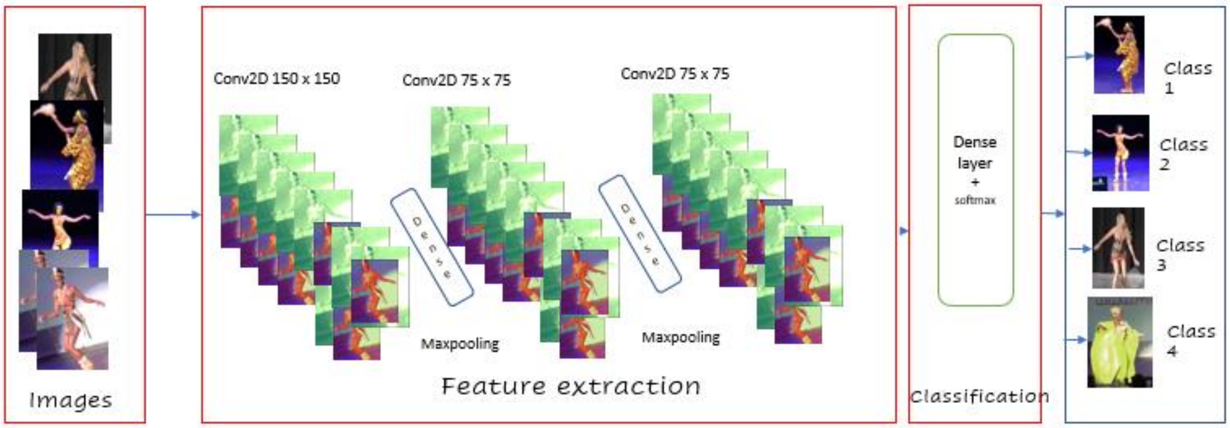


Figure 3.1: Dance Classification Model.

3.2.2 Dance modelling design

The dance modelling design was carried out using Pose estimation Algorithms. The pose estimation algorithms were used to identify the various body joints used in the dance and their positions in each pose. The dance modelling architecture (Figure 3.2) begins with splitting the dance video into different frames. The poses in each of these frames were identified using open pose and deep pose algorithms. The pose in each frame was exported into a dictionary in order to have the orientation of each of the body pose stored. The numbers were then used to plot the pose stick figure.

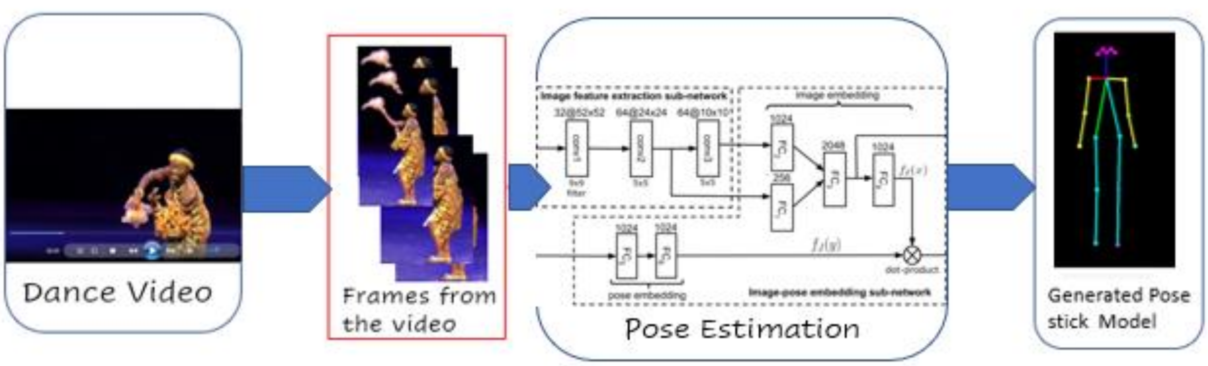


Figure 3.2: Dance Modelling Architecture

3.3 Data Collection

The dance dataset was collected from YouTube videos. YouTube has several dance videos which are freely available for use. Videos of several dancers were collected and grouped into classes. The videos were reduced to short and similar length to ensure evenness and make processing faster. A total of hundred videos were saved into the Traditional African Dances Dataset. For the classification process, three dances were zoomed in on. The decision to concentrate on these three dances were based on the knowledge of the researcher about them, number of videos available for the dance in the database and the distinct characteristics of those dances. The three dance classes are Adowa, Bata and Swange. Seven thousand seven hundred and ninety-one images were generated and used for the classification. Figure 3.3 presents some dance images.

Figure 3.3: Dance Data Collection

3.4 <u>Data Preparation</u>

The videos that belong to the dance classes that will be investigated were collected and processed for use. The first step in data processing was the clipping of the videos. The videos were clipped according to the parts that emphasizes the major poses in the dance. Frames (images) were extracted from the videos. The background from some of the images were extracted in order to reduce the number of pixels the algorithm is training on. After this Image augmentation algorithms were applied to the images. This is to introduce some noise into the dataset such that the algorithm is still able to predict right when the orientation of the images is not upright. The dataset was then divided into training and testing dataset at ratio 80:20.

3.5 Model evaluation

3.5.1 Dance Classification

The dance classification models were evaluated using the following metrics: Accuracy, Confusion matrix, precision, and recall.

Accuracy

This is the metric that measures how close the prediction made is close to the exact class that the dance belongs to. This will provide information about how well the model performed. Figure 3.4 is the formula for calculating accuracy.

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$

Figure 3.4 : Formula for calculating accuracy.

The accuracy of the model on the training and testing datasets will be compared.

Confusion matrix

The confusion matrix provides information about the number of true positive, false positives, true negatives, and false negatives. True positives are the images predicted as positives which are positive. False positives are the images predicted as belonging to a class but do not belong to that class. True negatives are the number of images predicted not to be in a class and do not belong to that class. False negatives are the number of images predicted not to be in a class, but they belong to that class. This is an improvement over accuracy.

Precision

This metric answers the question of the proportion of positive identifications that were correct. Another name for it is true positive rate. It is calculated using the following formula.

$$Precision = \frac{True\ positive}{True\ positive + False\ positive}$$

Recall

It is the fraction of the true positives that were identified. It is also referred to as sensitivity. The formula for calculating recall is as follows.

$$Recall = \frac{True\ positive}{True\ positive + False\ negative}$$

3.5.2 Dance modelling evaluation

The dance model was evaluated based on human evaluation of how close the pose stick generated is close to the dance video.

3.6 Conclusion

This section has provided information about the research questions that the study aims to answer, how the data will be collected and processed. It also provided information about the model architecture as well as how the model will be evaluated.

CHAPTER 4. ANALYSIS AND PRESENTATION OF FINDINGS

4.1 Introduction

Discussions and analysis of the data collection, experimentation, and findings of the study on classification, digitization, and preservation of Traditional African Dances are detailed in this chapter. The following research questions were investigated in the study.

a. Can YouTube videos be used to generate sufficient data for Traditional dance classification and modelling using deep learning techniques?

b. What deep learning technique(s) will identify and classify Traditional African dance poses?

c. Can deep learning techniques generate a 'tangible' dance model of Traditional African Dance from videos?

d. Are deep learning techniques suitable for developing a framework for intangible cultural preservation?

The hypothesis that were investigated includes:

a. YouTube videos will be enough to generate hundred thousand images for dance classification study.
b. Convolutional Neural Network (CNN) will classify Traditional African dance at 95% accuracy.
c. Deep learning techniques will generate a Traditional African dance pose stick model from videos from videos at 95% accuracy.

4.2 Research Question 1

Can YouTube videos be used to generate sufficient data for Traditional dance classification and modelling using deep learning techniques?

The first purpose of this study was to generate Traditional African Dance Dataset (TAD^2) because there is none in existence to the best of the researcher's knowledge. The dataset was generated following the TAD^2 design (Fig4.1).

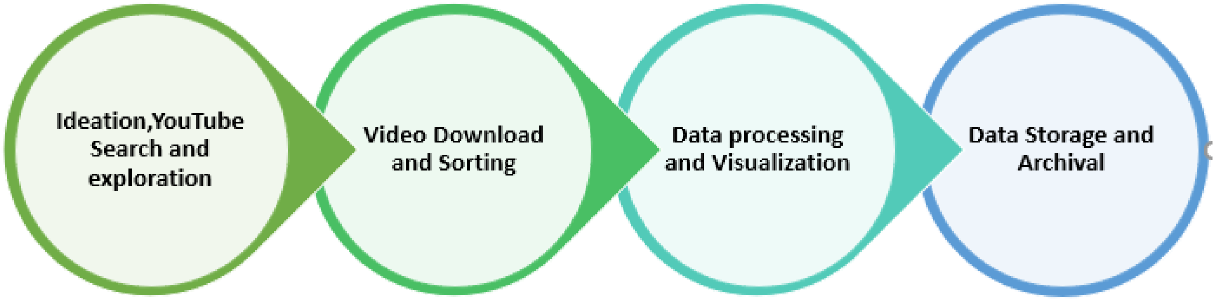


Figure 4.1: TAD^2 design process

4.2.1 Ideation, YouTube Search and Exploration

YouTube is a large collection of videos with five hundred (500) hours of video content uploaded per minute (Aslam, 2021). A list of popular Traditional African dance was put together. This list was used to make informed search of YouTube videos for the video collections that can be used for deep learning studies. Despite the large number of videos loaded, in most cases only few can be used because many of the videos were poorly recorded while some were wrongly labelled. The factors that determine whether a video should be downloaded includes, video length, dancer visibility, background, video quality and video content. The videos downloaded were restricted to a maximum of twenty-five minutes. Most videos are between One (1) and three (3) minutes. This is because videos of longer length are not the best for video processing. Longer videos also have other contents that were not important to us in most cases.

4.2.2 Video Download and Sorting

Videos that meet the requirements decided upon were downloaded using the 4kVideoDownloader (2020). The software provides the opportunity to download YouTube videos in four easy steps. Copy the video link from YouTube, paste the link in the designated tab in the software interface, choose the video quality of choice and click on the download button. The videos were then sorted into the different classes of dance. A total of hundred videos of various dances were downloaded.

4.2.3 Data Processing and Visualization

Most of the videos were converted to images so that the images can be used for study. This is the easiest format to format videos knowing that a video is a collection of images. A total of twelve thousand images were generated from some of the videos. Some videos are left as they were for future processing for other studies. The images generated were annotated Tzutalin LabelImg (2015). LabelImg is an image annotation software which export image annotations as Xml files. The images were saved in folders of the respective classes. The dance classes are Sinte, Adowa, Bata, Swange, Guinea, Botswana and South Africa. The datasso consists of audio recordings of the Traditional Dance drums. The audio recordings were extracted from videos of the Traditional Dance performance on YouTube. The downloaded audios belong to only three classes- Bata, Adowa and Swange. Table 4.1 provides information about the Video, image and audio sound dataset.

The image dataset was further explored for visualization purpose. 7751 images belonging to three dance classes (Adowa, Bata, and Swange) were used for the visualization. Attention was given to these dances because they have the most balanced dataset. Figure 4.2 is the histogram of the dance visualization.

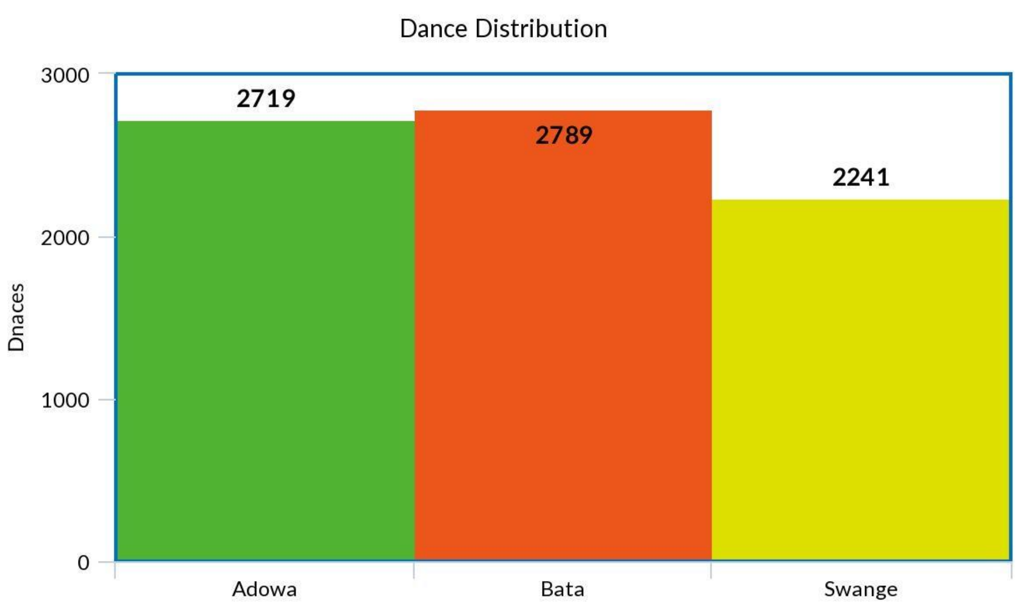
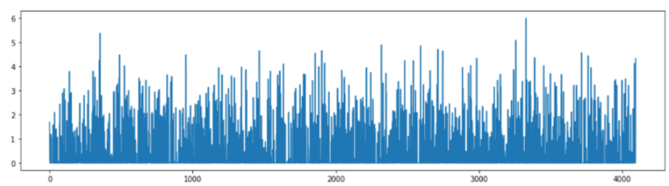


Figure 4.2 : Dance Dataset distribution for three classes.

t-Distributed Stochastic Neighbor Embedding (tSNE) visualization of the dataset was carried out on the unlabeled images. The images were analyzed using a pre-trained Convolutional Network, in this case VGG16. The network was used to extract feature vectors which represent each image. This process is unsupervised because the images are put into clusters mainly on the features extracted using VGG16 (Model summary in Appendix 1). The program was written in python using the following libraries: OS, Keras, Numpy and Matplotlib libraries. The Matplotlib library was important for the plotting of images and general visualization. First the image features were reduced to it array, which is further represented in an image; an example of this is shown in Figure 4.3.

a. Image



b. Feature plot

Figure 4.3: Image and feature plot of the image

After this is obtained the principal component of each image is generated using Principal Component Analysis. The principal component of the images obtained were used to identify similar images. This process uses the feature extracted from the images. It can use the features to identify images that are closely related. Since this is unsupervised, it only uses the information it acquired form the data to decide closely related images. Figure 4.4 is a visualization of such grouping of images using their principal components.

Figure 4.4: Grouping of images using their principal components.

The extracted features were finally used to generate the tSNE visualization of all the images. The generated visualization is a metric for evaluating our dataset. Knowing that without supervision the features were brilliantly used to group images from the same dance together. Figure 4.5 provides the visualization of the dataset.

Figure 4.5: tSNE visualization of TAD[2].

The tSNE visualization of the images provided a very important information about the close relation of the of some of the dance information. Some video sources were much larger than the others. Although they belong to the same classes the algorithm could map them to their different source. The tSNE visualization generated clusters using the principal features plotted from the principal components in the images.

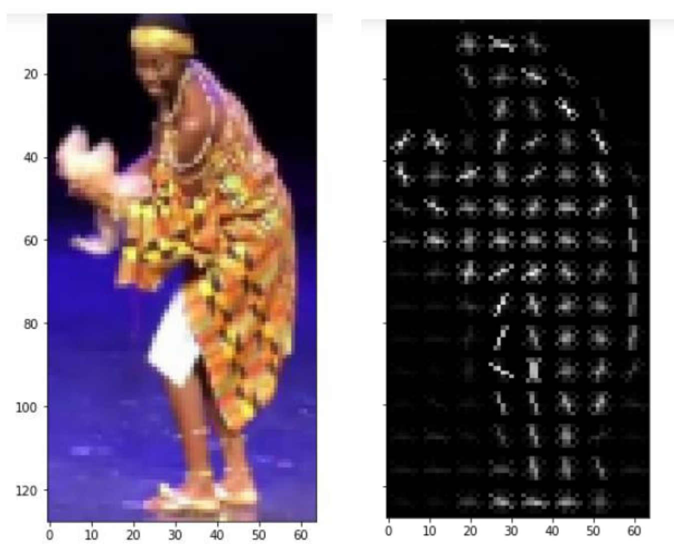
4.2.4 Data Storage and Archival

The video, images and audio data collected are all stored in an external hard disk in a secured and protected space. It is hoped that soon the data will be hosted on the cloud so it can be assessed by others to be used for related studies.

4.3 Research Question 2

What deep learning technique(s) will identify and classify Traditional African dance poses?

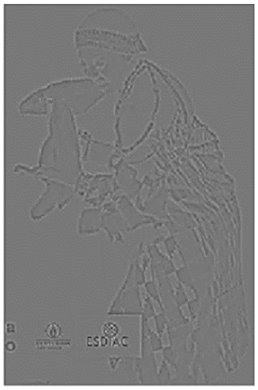
The convolutional Neural Network algorithms were used for the image classification process. Four different CNN based models were used for the classification process and each of them will be discussed. Before the image classification itself, the images were preprocessed. First the Histogram of Gradient (HOG) features of the images were extracted. Figure 4.6 provides the HOG features of one of the images in the dataset.



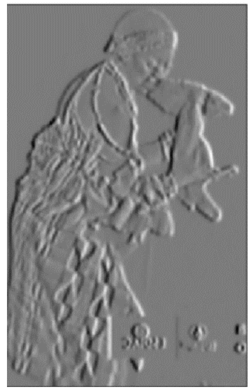
(a) (b)

Figure 4.6: HOG feature of one of the images.

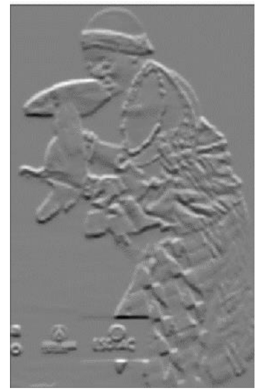
The HOG identifies important parts of the image. The HOG is generated from a combination of three different coordinates of the image. Figure 4.7 presents the HOG features of the image.



(a) Laplacian



(b) Sobel-x

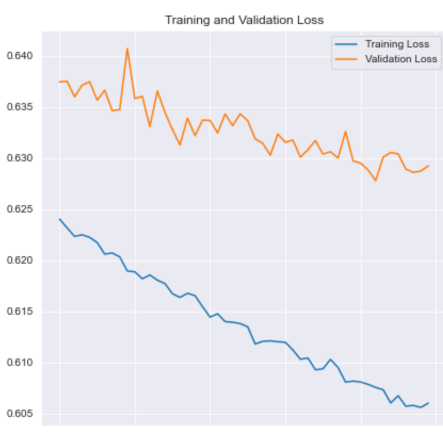


(c) Sobel-y

Figure 4.7 HOG components of the image.

4.3.1 MobileNet Classification

MobileNet is one of the benchmark classifications models that are being used to validate another model. MobileNet classification model is a convolution-based neural network that incorporates both the standard convolution layers and depth-wise convolution layers. An accuracy of 93% was achieved with a loss value of less 0.64 for both the training and validation set. (Figure 4.8). Table 4.2 provides the summary of the precision, recall and F1 score for the individual dance classes and the average accuracy.



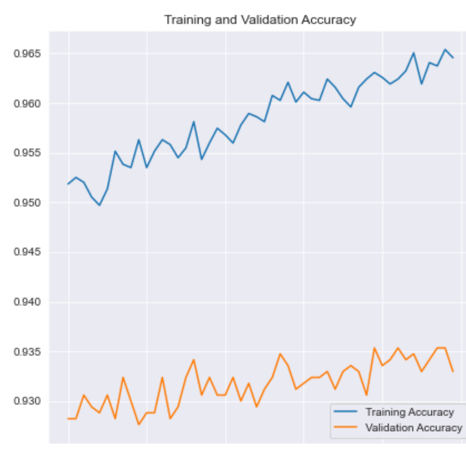


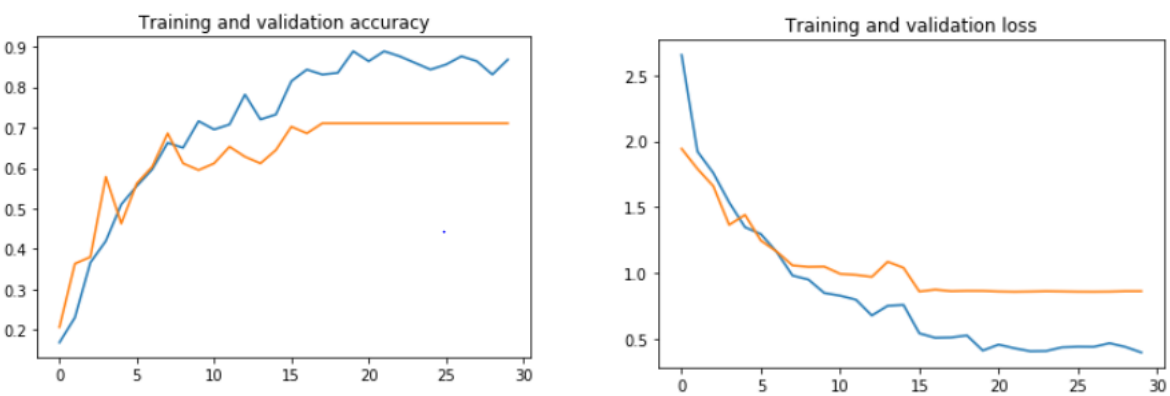
Figure 4.8: (a)Training and validation loss and (b) Training and Validation accuracy for the dance dataset.

Table 4.1: Results of the Dance dataset using MobileNet

	Precision	Recall	F1 Score
Adowa (Class 1)	0.95	0.94	0.95
Bata (Class 2)	0.98	0.95	0.97
Swange (Class 3)	0.73	0.85	0.78
Macro Average	0.89	0.91	0.90
Weighted Average	0.94	0.93	0.93
Accuracy	0.93		

4.3.2 RESNET50

Another benchmark image classification model that was used is RESNET50. The RESNET50 model was used to classify the images and an accuracy of 93% was achieved with a loss value of less 0.64 for both the training and validation set. (Figure 4.9). Table 4.3 provides the summary of the precision, recall and F1 score for the individual dance classes and the average accuracy.



(a) (b)

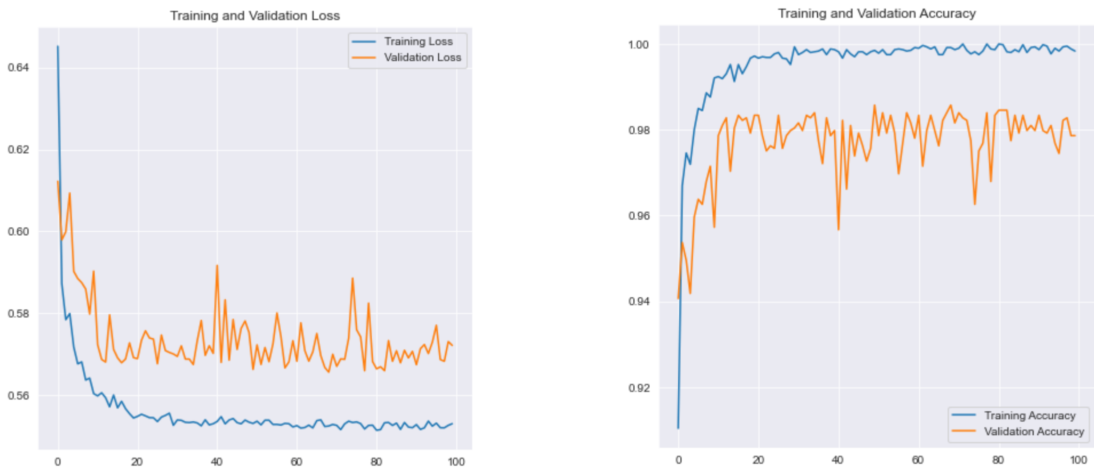
Figure 4.9: (a) Training and validation loss and (b) Training and Validation accuracy for the dance dataset.

Table 4.2: Results of the Dance dataset using RESNET50

	Precision	**Recall**	**F1 Score**
Adowa (Class 1)	0.99	0.95	0.98
Bata (Class 2)	0.98	0.96	0.97
Swange (Class 3)	0.92	0.95	0.94
Macro Average	0.92	0.91	0.92
Weighted Average	0.98	0.98	0.99
Accuracy	0.98		

4.3.1 MobilenetV2

Mobilenet V2 was also used for the classification. However, four additional layers (One maxpooling, one dropout and two dense) were added to the model architecture. This led to an improvement on the performance of the model. An accuracy of 98% was achieved with a loss value of less 0.62 for both the training and validation set. (Figure 4.10). Table 4.4 provides the summary of the precision, recall and F1 score for the individual dance classes and the average accuracy.



(a) (b)

Figure 4.10: (a)Training and validation loss and (b) Training and Validation accuracy for the dance dataset.

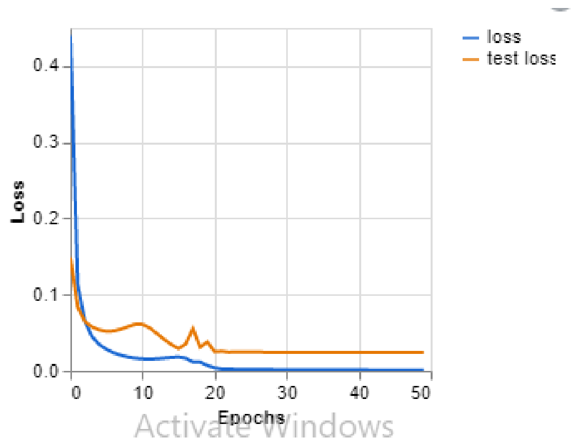
Table 4.3: Results of the Dance dataset using MobilenetV2

	Precision	**Recall**	**F1 Score**
Adowa (Class 1)	0.98	0.98	0.98
Bata (Class 2)	0.99	0.98	0.99
Swange (Class 3)	0.95	0.94	0.95
Macro Average	0.97	0.97	0.97
Weighted Average	0.98	0.98	0.98
Accuracy	0.98		

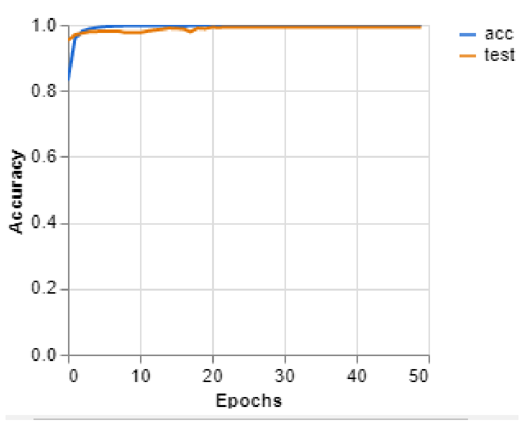
With this model, not only did the average accuracy improved, it was also an improvement in every dance class.

4.3.2 TAD^2 Model 1

An image classification model using MobileNet as a base model was used. Two additional Conve2D layers were added to it. An accuracy of 99% was achieved with a loss value of less 0.4 for both the training and validation set. (Figure 4.11).



(a)



(b)

Figure 4.11: (a)Training and validation loss and (b) Training and Validation accuracy for the dance dataset.

4.3.3 TAD^2 Model 2

A custom-built model was developed for the image dataset. Figure 4.12 provides the model architecture.

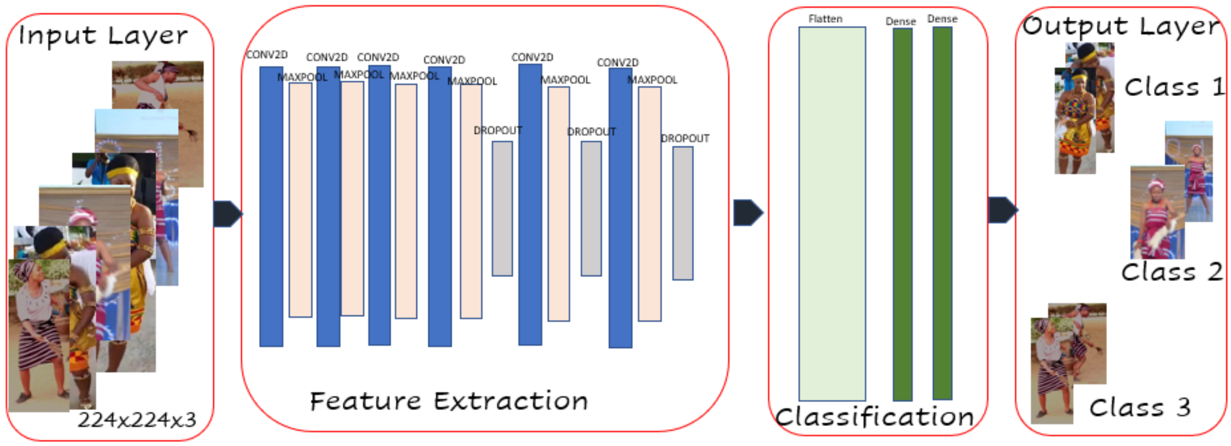


Figure 4.12: Model Architecture

This model has a smaller number of layers than most of the benchmark models. There was also no transfer of learning (no base model was used). In the input layer images are fed in. All the images were given a uniform shape of 224x224x3. The dance classification was performed using all the color channels. There are to major important stages in the classification process.

Stage 1: Feature extraction

These layers were used to extract the feature maps of the images. The CONV2D layer performed convolution on the image vectors. The mathematics behind the is presented by equation 1. The first two layers are for the extraction of the low-level features for example lines, edges, and corners. Figure 4.13 is a visualization of some feature map extraction performed in the convolution layers.

$$G(x,y) = \omega * F(x,y) = \sum_{\delta x=-k_i}^{k_i} \sum_{\delta y=-k_j}^{k_j} \omega(\delta x, \delta y) \cdot F(x+\delta x, y+\delta y)$$

where ω is a kernel and $-k_i \leq \delta x \leq k_i, -k_j \leq \delta y \leq k_j$ its elements.

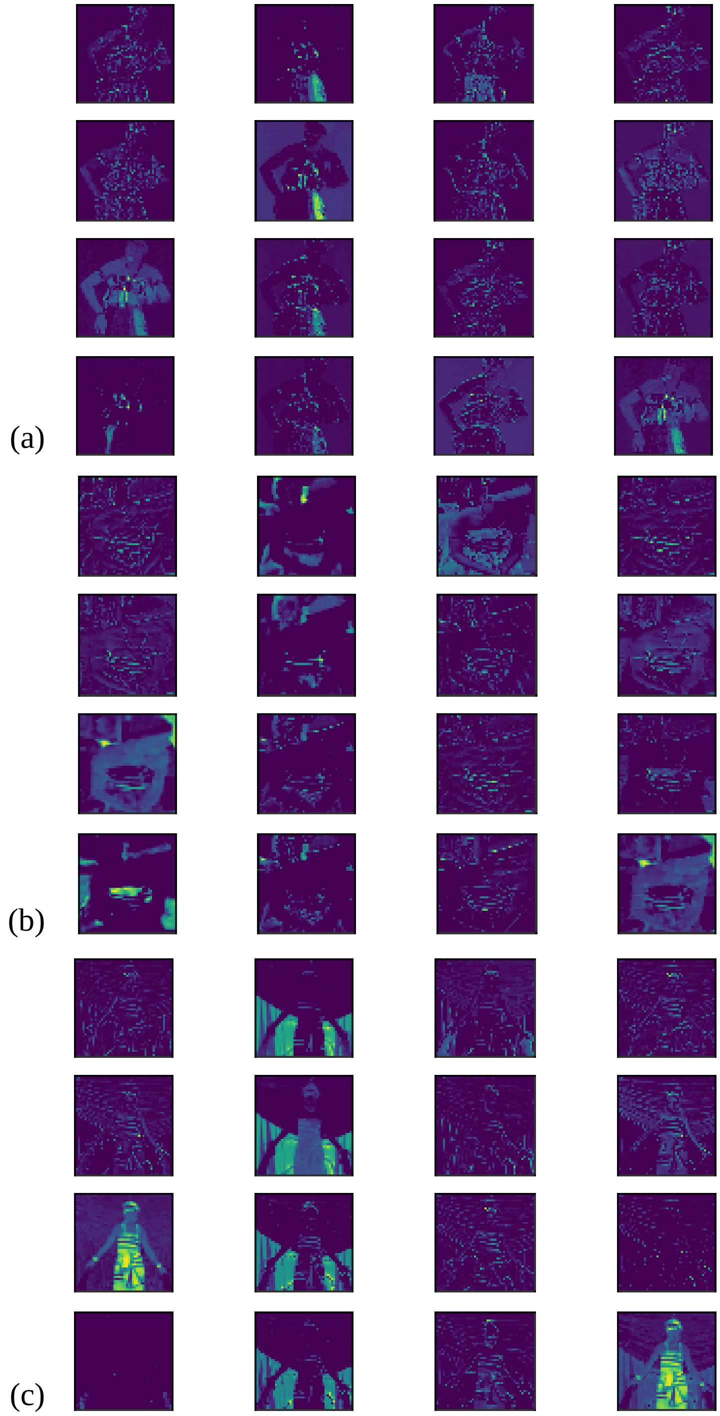
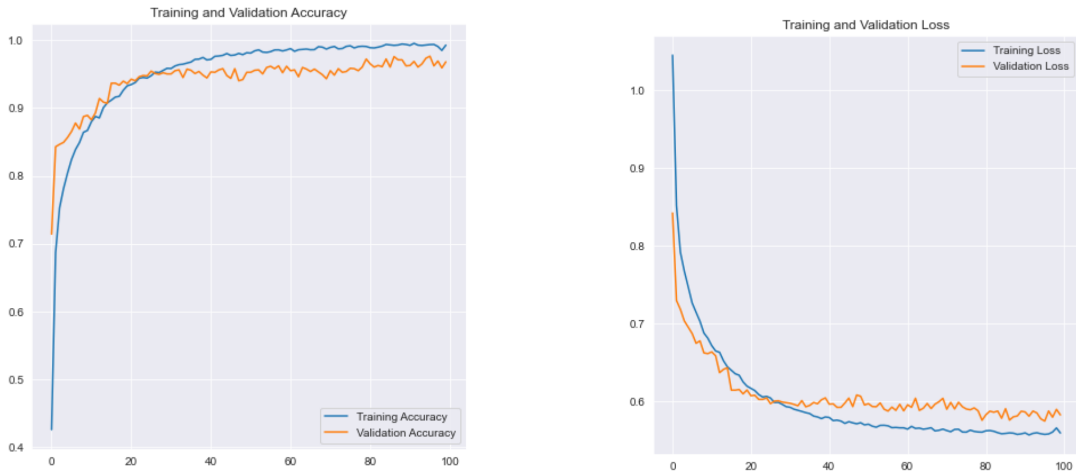


Figure 4.13: Feature map visualization of images from each dance class.

Stage 2: Classification

In the second stage the proper classification effort takes place. The flatten layer converts the data into a one-dimensional array. This is the same as what was shown in Figure3b. The one-dimensional array (a single long list of feature vector) from the flattening is then fed into the dense layer. The dense layer is also known as the fully connected layer. This layer connects all the neurons fed in from the flatten layer. This process eventually leads to the classification. An accuracy of 97% was achieved with a loss value of less 0.62 for both the training and validation set. (Figure 4.14). Table 4.6 provides the summary of the precision, recall and F1 score for the individual dance classes and the average accuracy.



(a) (b)

Figure 4.14: (a)Training and validation accuracy and (b) Training and Validation loss for the dance dataset.

Table 4.4: Results of the Dance dataset using TAD^2 model 2

	Precision	**Recall**	**F1 Score**
Adowa (Class 1)	0.96	0.99	0.97
Bata (Class 2)	0.99	0.98	0.98
Swange (Class 3)	0.93	0.87	0.90
Macro Average	0.96	0.94	0.95
Weighted Average	0.97	0.97	0.97
Accuracy	0.97		

From the result we can conclude that a model with fewer number of layers, less time for training and fewer resources consumption performed as well as benchmark models with a lot more requirements.

4.3.4 Sound Classification

Dataset

The data set for the sound classification was put together using the recordings of the drumbeats that are performed for the dance. This sound classification was carried out to enhance the classification process. The dataset is a collection of .wav files of the sound. The sound record is an ensemble of all the instruments used for the dance performance. In the case of Bata dance the bata drums are the major instruments used. The bata ensemble is made up of three drums. Two of the drums are larger and are called Iya_Ilu and Omele-Abo; both are double-headed. The third drum is smaller, and it is called Omele-Ako. Omele-Ako is a combination of three very small drums tied together and are played by a single person. Omele-Ako is the speaking drum among the three and it usually dictates the dancer's actions. For the Adowa dance, the ensemble consists of Atumpan, Astimeyu, two metal bells called Dawuro, gourd rattles (Ntrowa) and four smaller drums. The Swange dance music is mostly dominated by Khakaki (A type of trumpet), and Gbande drums.

Data Exploration

The librosa library was used to load the sound files and the wave plots of the music was produced. The wave plot of the sounds provides the visual representation of the sound. The wave plot of two files from each of the dance classes was produced. Figure 4.15 provide a visualization of some of the sounds in each class. The wave plots showed that there are variations in the sound produced, this is especially dependent on the sound clip, stage of dance, drummers, and many other factors. The linear spectrogram of sounds (figure 4:16) in each class did not also provide us with distinct differences that can be used to tell one class from the other using the images.

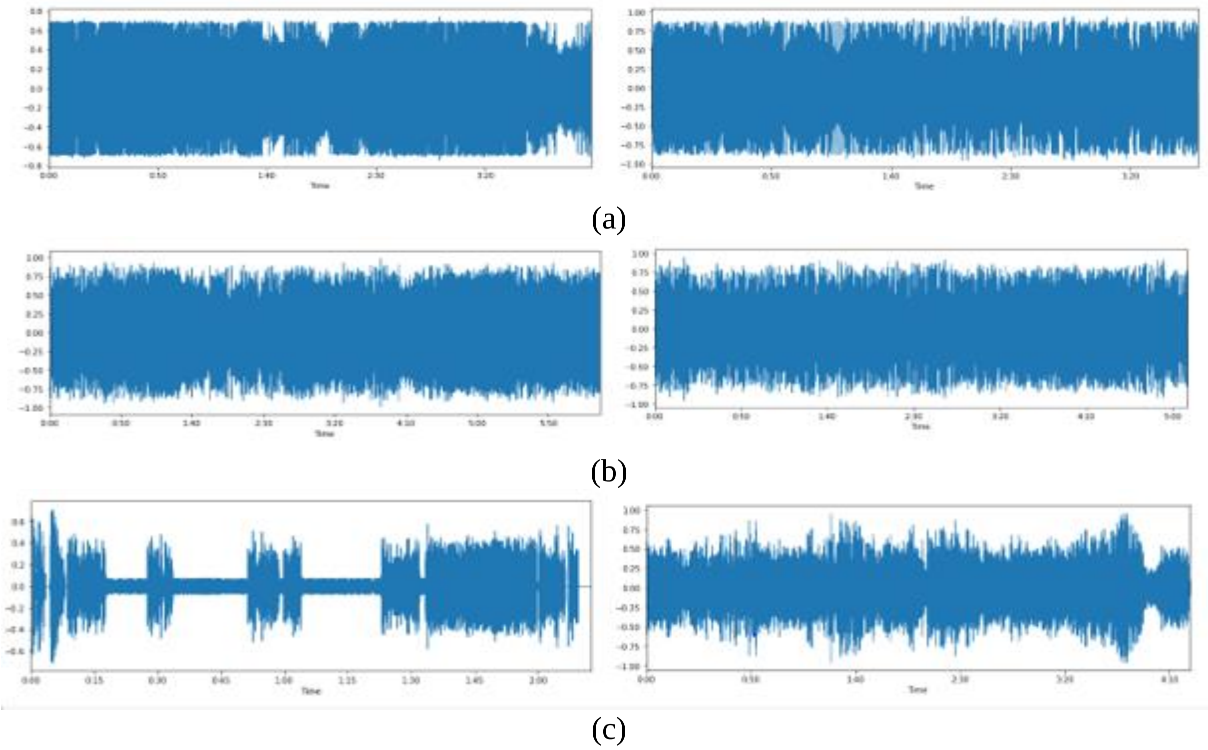


Figure 4.15: (a) Adowa (b) Bata (c) Swange wave plots.

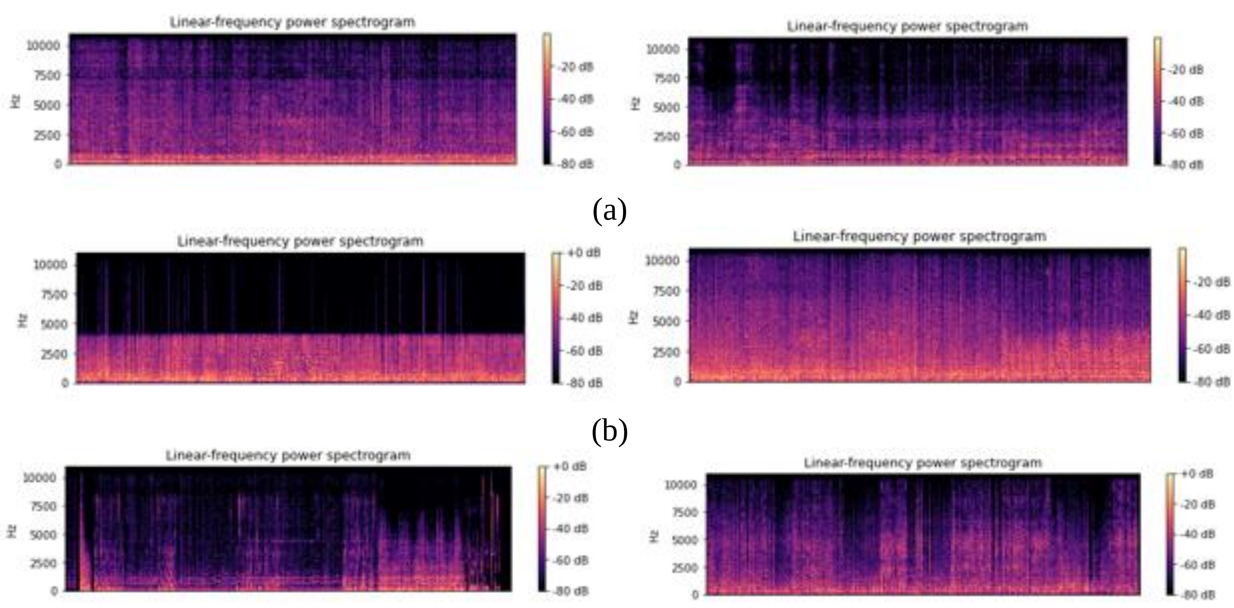


Figure 4.16: (a) Adowa (b) Bata (c) Swange Linear Frequency Power Spectrogram

Feature Extraction

The features of the sound were extracted using the Mel-Frequency Cepstral Coefficients (MFCC) technique. This technique uses a quasi-logarithmic spaced frequency scale rather than a linear spaced frequency scale used by the spectrogram. The MFCC technique is very close to how sound is processed by human beings.

Classification

The model for the classification process was based on CNN (Figure 4.16).

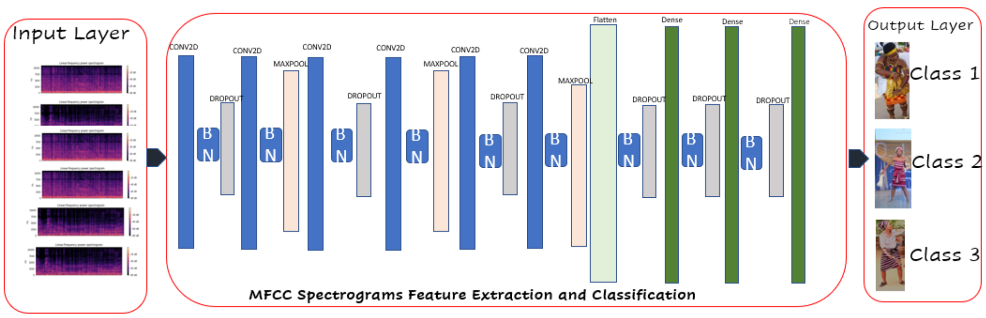
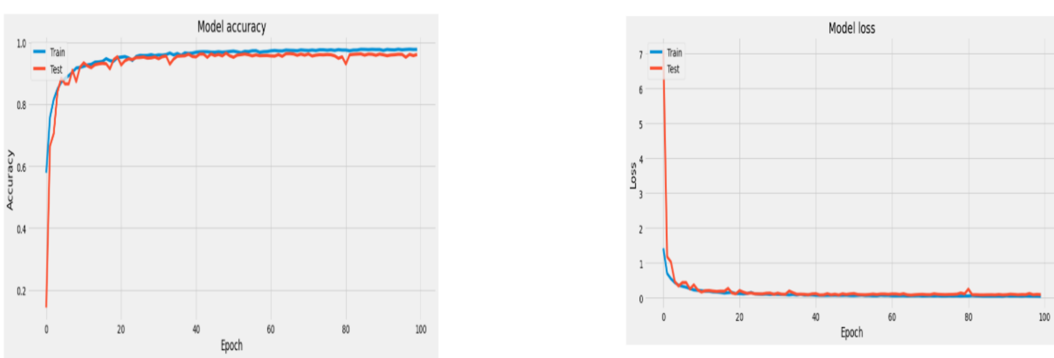


Figure 4.17: TAD^2 Sound Classification Model Architecture

The training of the model was carried out using different learning rates, activation functions and number of epochs. The best model was archived using a learning rate of 0.001, Stochastic Gradient Descent (SGD), Adam optimizer and twenty-five epochs. An accuracy of 96% was achieved with a loss value of less 0.2 for both the training and validation set. (Figure 4.18). Table 4.7 provides the summary of the precision, recall and F1 score for the individual dance classes and the average accuracy.



(a) (b)

Figure 4.18: (a)Training and validation accuracy and (b) Training and Validation loss for the dance dataset.

Table 4.5: Results of the Dance dataset using TAD^2 model 3

	Precision	**Recall**	**F1 Score**
Adowa (Class 1)	0.99	0.99	0.98
Bata (Class 2)	0.99	0.98	0.98
Swange (Class 3)	0.98	0.99	0.99
Macro Average	0.93	0.94	0.93
Weighted Average	0.96	0.96	0.96
Accuracy	0.96		

The model was efficient in the sound classification process.

4.4 Research Question 3

Can deep learning techniques generate a 'tangible' pose stick dance model of Traditional African Dance from videos?

4.4.1 Pose Estimation

Human action documentation is usually performed using data collected from Kinect sensors. This has faced a lot of challenges. It was the first real-time pose estimation that was able to estimate human poses from videos. The dance modelling architecture is presented in Figure 4.19.

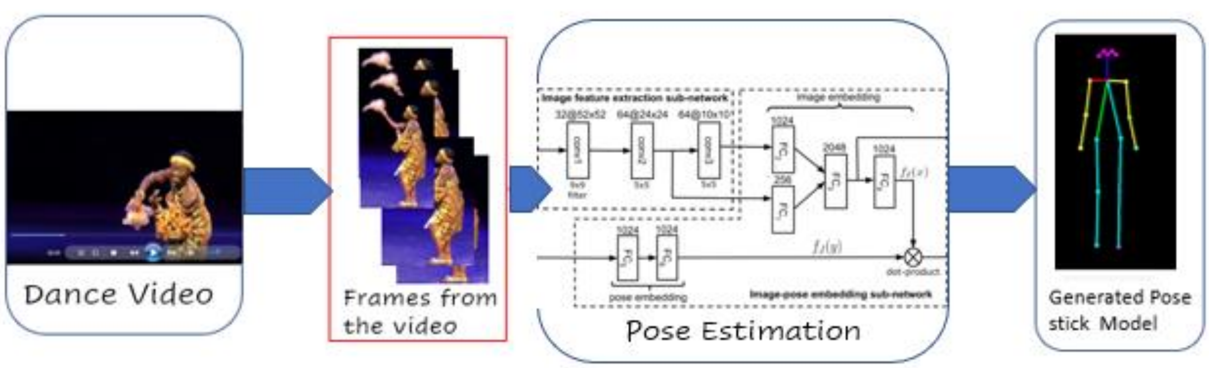


Figure 4.19: Dance Pose stick Model Generation Architecture.

The Open pose algorithm (Cao, Hidalgo, Simon, Wei and Sheikh, 2019) was developed by the researchers in Carnegie Mellon University. The Google deep pose algorithm (Toshev & Szegedy, 2014) was also applied to the dance dataset. Figure 4.20 and Figure 4.21 presents images from the dance pose estimation using Open pose and deep pose respectively.

(a)

(b)

(c)

(d)

Figure 4.20: Adowa (b and c) and Sinte (a and d) pose estimation from dance video using Open pose algorithm.

(a)

(b)

(c)

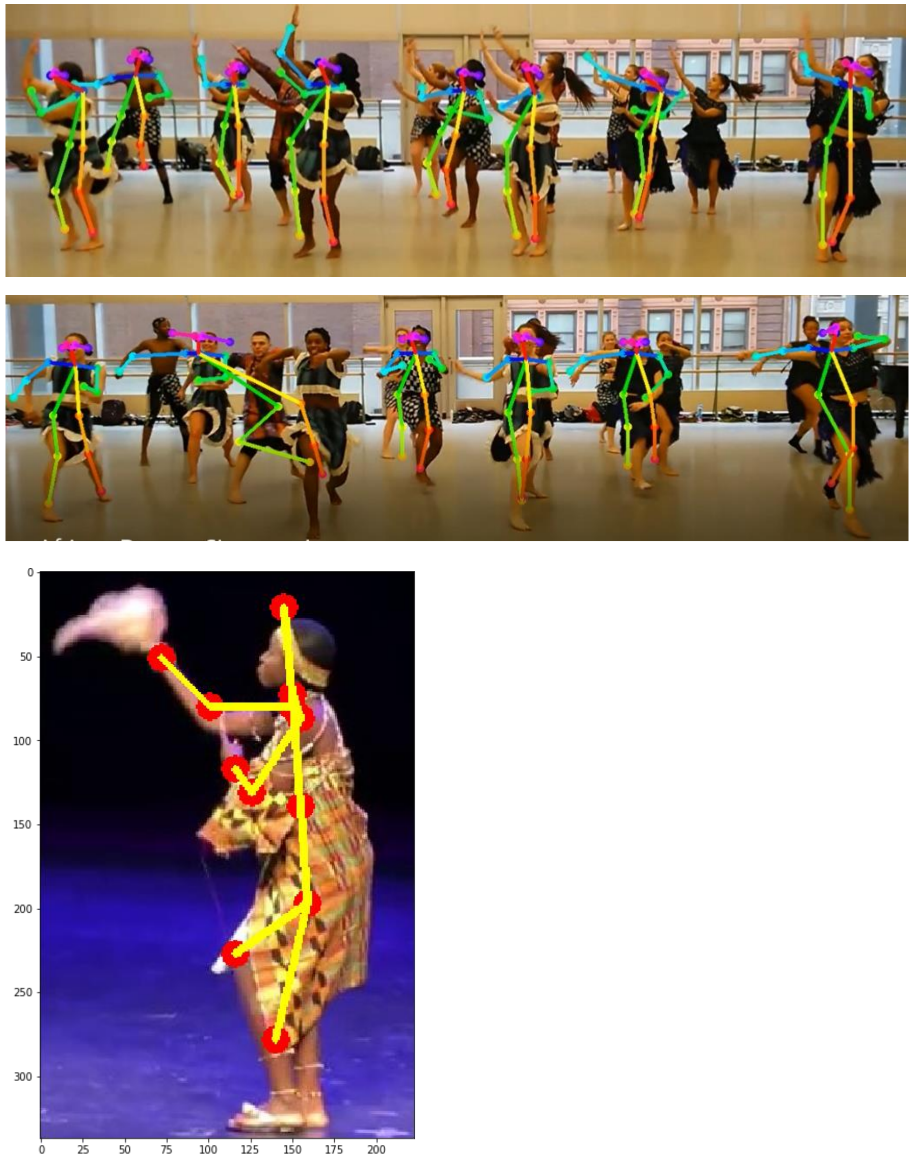


Figure 4.21: Sinte (a and b) and Adowa (c) pose estimation from dance video using deep pose algorithm.

From the poses generated from both algorithms, the CMU pose estimation produced the best pose estimation result. Something that is clear from both pose estimations are that body part occlusion with clothes limited the effectiveness of the pose estimation. Figure 4.20c proves these. The long garment occluded the algorithm from being able to estimate the poses appropriately. The CMU pose estimation was leveraged upon to generate the pose stick.

4.4.2 Pose Stick Generation

The pose estimation algorithm was used to capture the location of each body joints per frame and the value were exported into a dictionary.

{'bp_0_x': 0.6296296296296297,'bp_0_y': 0.842391304347826,'bp_0_score': 0.9075720906257629,
'bp_1_x': 0.6388888888888888,'bp_1_y': 0.7717391304347826,'bp_1_score': 0.9162707328796387,
'bp_2_x': 0.6018518518518519,'bp_2_y': 0.7717391304347826,'bp_2_score': 0.834660530090332,
'bp_3_x': 0.5416666666666666,'bp_3_y': 0.7608695652173914,'bp_3_score': 0.7607961893081665,
'bp_4_x': 0.5046296296296297,'bp_4_y': 0.8206521739130435,'bp_4_score': 0.6363097429275513,
'bp_5_x': 0.6759259259259259,'bp_5_y': 0.7717391304347826,'bp_5_score': 0.9108472466468811,
'bp_6_x': 0.7361111111111112,'bp_6_y': 0.7228260869565217,'bp_6_score': 0.7615177035331726,
'bp_7_x': 0.8101851851851852,'bp_7_y': 0.7065217391304348,'bp_7_score': 0.6132577657699585,
'bp_8_x': 0.6018518518518519,'bp_8_y': 0.5923913043478262,'bp_8_score': 0.7169932126998901,
'bp_9_x': 0.5277777777777778,'bp_9_y': 0.6141304347826086,'bp_9_score': 0.7349910736083984,
'bp_10_x': 0.5324074074074074,'bp_10_y': 0.47282608695652173,'bp_10_score': 0.7490565180778503,
'bp_11_x': 0.6481481481481481,'bp_11_y': 0.5760869565217391,'bp_11_score': 0.6652555465698242,
'bp_12_x': 0.6574074074074074,'bp_12_y': 0.4130434782608695,'bp_12_score': 0.7154346704483032,
'bp_13_x': 0.6712962962962963,'bp_13_y': 0.27717391304347827,'bp_13_score': 0.8217236995697021,
'bp_14_x': 0.6203703703703703,'bp_14_y': 0.8532608695652174,'bp_14_score': 0.9057081937789917,
'bp_15_x': 0.6388888888888888,'bp_15_y': 0.8532608695652174,'bp_15_score': 0.8873612880706787,
'bp_16_x': 0.6111111111111112,'bp_16_y': 0.842391304347826,'bp_16_score': 0.5647059082984924,
'bp_17_x': 0.6574074074074074,'bp_17_y': 0.8478260869565217,'bp_17_score': 0.7238561511039734}

The dictionary contains information about the x and y coordinate of 17 distinct body joints and provides the confidence score for the positioning. These values are collated together in the video and used to generate the pose stick model. The pose sticks generated can be compared to dance labanotation. However, this has more granularity. Only the Sinte dance was used for this layer of experimentation. These collated values were exported into data frames (Figure 4.22) which was then used to generate a dance graph (Figure 4.23).

	bp_0_score	bp_0_x	bp_0_y	bp_10_score	bp_10_x	bp_10_y	bp_11_score	bp_11_x	bp_11_y	bp_12_score	bp_12_x
count	2965.000000	2965.000000	2965.000000	3086.000000	3086.000000	3086.000000	3102.000000	3102.000000	3102.000000	3097.000000	3097.000000
mean	0.886319	0.563867	0.813445	0.764526	0.496623	0.287827	0.713336	0.584999	0.547711	0.760105	0.610465
std	0.103659	0.052015	0.032561	0.117346	0.058229	0.040618	0.066374	0.054805	0.026035	0.103240	0.062118
min	0.117926	0.444444	0.500000	0.096528	0.328704	0.027174	0.175995	0.240741	0.380435	0.088589	0.254630
25%	0.848838	0.523148	0.793478	0.737303	0.462963	0.271739	0.681005	0.541667	0.527174	0.724404	0.560185
50%	0.897846	0.555556	0.809783	0.789244	0.504630	0.282609	0.719785	0.578704	0.543478	0.773168	0.611111
75%	0.949189	0.601852	0.831522	0.836083	0.532407	0.288043	0.757061	0.629630	0.565217	0.821024	0.666667
max	1.053323	0.708333	0.983696	0.929204	0.708333	0.510870	0.869126	0.740741	0.679348	0.951203	0.773148

Figure 4.22: Description of the exported data frame from the pose estimation.

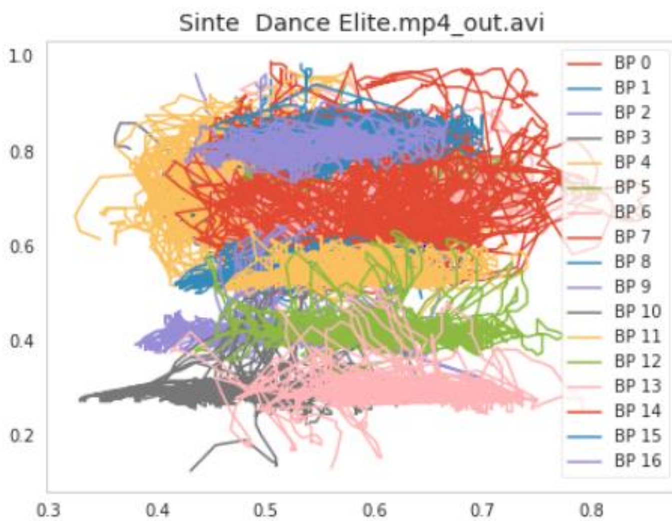


Figure 4.23: Sinte Dance Graph generated from the pose estimation from video.

The pose sticks generated form the pose estimation was exported frame after frame. These poses were then merged in a gif file. Figure 4.24 presents the dance thumbnails of the dance sequence before they were made into a video.

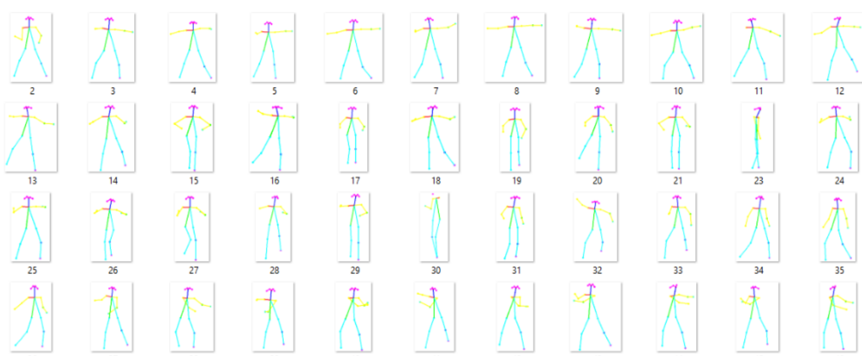


Figure 4.24: Dance pose sticks generated in sequence

4.5 Research Question 4

Are deep learning techniques suitable for developing a framework for intangible cultural preservation?

The process involved in the design of this study was made into a framework that can be used for intangible cultural heritage preservation and learning. The Deep Culture Preservation

Framework presents a framework that can be adapted to most intangible cultural heritages. The framework has three main parts (Figure 4.25).

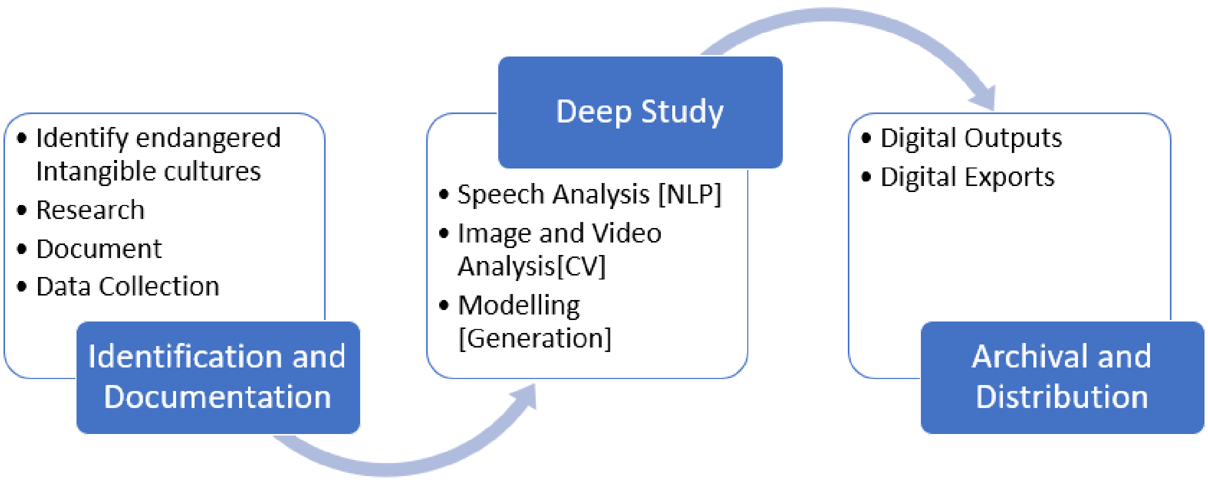


Figure 4.25: Deep Culture Preservation Framework

4.5.1 Stage A: Identification and documentation

This stage is subdivided into four processes. These processes are continuous running iteratively.

Figure 4.26: Identification and Documentation Process.

Identify Endangered Intangible Cultures

This step focusses on getting information about the cultures that are fading aware. The UNESCO list of such cultures is a good way to start.

Research

This step is about researching into the identified cultures. Often the surface information is always a small tip of the iceberg. There will be a lot to research about the identified culture, its importance, elements, history, and everything about the culture. Sometimes, a research into an intangible heritage can lead to another that needed earlier rescue. This process itself is continuous and can go for a long time. The researcher should be able to identify when enough information for moving forward has been reached.

Document

It is important to document all findings. The more information available, the more one can make informed decisions. Every detail should be put in an organized record. No information is too much to document.

Data Collection

This stage will be informed by the research that has been carried out. What are the important data to be collected to digitize the cultural heritage for preservation? The data can be videos, images, materials, language/text, process souvenirs etc.

4.5.2 Deep Study

After the important data for the cultural preservation has been collected, the next thing is to perform a deep study on the culture as well as data. Assuming that a culture to be preserved is a song; for example, ‘Ekun Iyawo’, a song performed by brides on their last night under parents’ roof in the Yoruba Culture of Nigeria. Because this is a song, the data to be collected may be sound recording of the song or the text of the song. If the song is recorded as a text, the text data can be analyzed using Natural language processing deep learning techniques. In case data is a song, the sound waves can be used for the digitization process. There are various deep learning techniques available for different tasks.

4.5.3 Archival and Exports

There are two important pieces to cultural preservation which are storage and making it available to people. It is not enough to have important pieces of art stacked away behind some shelfs as a way of keeping them safe, the pieces of art should as well be made available for people to access and learn from. Culture is preserved when it is passed on from one generation to another. Whatever the digitized version derived from the dee study stage, it must be kept save (archived) and made available for cultural learning and preservation. For example, in the current study, pose stick models of dance were archived but also presented in a data format that can be used in order software for development. For example, the 2d pose stick model can be textured and exported into Unity for a game development effort. The same dance model cane be used in animated movies etc. Making the digitized model available opens the culture to different population who can use it effectively for common good.

CHAPTER 5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusion and Recommendations

The conclusions and recommendations that were made from this study are detailed in this chapter. The study was carried out to develop an image and video database of Traditional African dances that could be used for action recognition and prediction studies as well as archived for cultural preservation purposes. Development of a deep learning algorithm for the classification of Traditional African dance was also investigated during this study. Another important piece in the study is the investigation of how deep learning algorithms can be used for the modelling of a pose stick dancer that can be used for dance preservation. The experiences gathered in the process was then used to develop a framework for intangible cultural preservation using deep learning techniques.

5.1.1 Traditional African Dance Dataset

The study revealed that sufficient information for a Traditional African dance study can be generated from YouTube. A very important thing to note is that information generation from YouTube is largely dependent on the ability of the researcher to use particular keywords for the search. The quality of the information that will be generated is determined by the expertise of the researcher to use specific terms to make the search a success.

YouTube is a goldmine for collaborative data collection especially for a video study. This is in agreement with the efforts of some other researchers who used Youtube for data collection. Abu-El-Haija, Kothari, Lee, Natsev, Toderici, Varadarajan and Vijayanarasimhan (2016) developed a dataset called YouTube 8M. The data set is the largest collection of multi-label video classification dataset, which consists of about 8 million videos. The dataset is annotated for computer vision studies. Kishore, Kumar and Kumar (2018) also used some YouTube videos in combination with their recorded video for the Indian Classical dance classification efforts. A major challenge with using uploaded videos on YouTube is that it is very error prone. Any YouTuber can upload a video and give it any title he or she chooses. In most cases people upload videos using titles that they believe will promote the viewing of their channels. This sometimes led to misinformation by such people.

Thus, it is important that any researcher embarking on using YouTube videos should dive deep into the knowledge of what the content of their video choice should be. In the case of Traditional African dance, it was important that the researcher possess good knowledge about the distinction of various African dances from literature or interviews before embarking on such a study. Another important caution on the part of the researcher is ensuring that all videos collected are viewed before it is being stored in the database. Some videos have contents that may be deviations from what is intended. The researcher should also ensure that irrelevant portions or deviations from intended contents are removed.

Video clipping is another important skill in this study. Several of the videos downloaded are from events where the dance is performed. Thus, it is important to reduce the video to what is important in the study rather than load all the videos in the database. Loading full lengths of activities that are not necessary will occupy unnecessary space in the database memory.

5.1.2 Traditional African Dance Classification

Two Convolutional Neural Network models were developed for the classification of Traditional African Dance images and both achieved a great result. The dance classification was successful with accuracies greater than 95%. The TAD^2 model 1 was built on the MobileNetV2 weights. The second model was built without using any base model. It had conv2D layers sandwiched with dropout layers at intervals. This classification model also performed well. Thus, we can conclude that deep learning techniques are effective for Traditional African dance classification. This agrees with Kishore, Kumar, Kumar, Sastry, Kiran, Kumar, and Prasad (2018) who also classified Indian Classical dance using deep learning techniques. Although the Indian Classical dances have features that differ from African dances, deep learning remains an effective technique. Gupta (2020) also used RESNET 152, a deep learning technique for the classification of Indian dance. He achieved an accuracy of 94% when CNN was used. From all this studies, deep learning remains a reliable dance classification algorithm.

Another conclusion from this study is that transfer learning can improve classification model performance. This is observed from the performance of TAD^2 model 1, which was built on MobileNet CNN model. Although, the TAD^2 model 2 performed well, the first one had a shorter training time. Using the weights from the already trained model, is believed to have contributed to the speed and the better performance of the model. The 1% difference in the accuracy is a lot when considered from the point of view of scalability. This conclusion agrees with Zhuang, Qi, Duan, Xi, Zhu, Zhu, Xiong, and He (2020). They proved that the use of pre-trained weight saves time and generate better results in their study.

The study also showed that the effectiveness of a CNN model is not absolutely dependent on the number of layers. For large volumes of data and from benchmark dataset, the accuracy has been shown to improve with an added convolutional layer. This is evident in some benchmark models like RESNET that went from having eighteen (18), to thirty-four (34), fifty (50), one hundred and one (101) and now one hundred and fifty-two (152) layers. While literature has proved that this model improved, it is not necessarily right to conclude that the success of a model is absolutely about the number of layers. There are other factors like the learning rate, activation function, batch numbers and others which contributed to the achieved success in the performance of the model. TAD^2 model 2 had very few layers but performed as well as the other CNN models used.

5.1.3 Traditional African Dance Sound Classification

The classification of the drumbeats for the three dance classes was carried out and a high accuracy and precision value was achieved. The drumbeats generated different waves that were distinguishable by the model. The sound classification model establishes the fact that the sound waves could be manipulated for music generation. The African drum has the reputation for being able to talk. All the drum ensembles had a leading drum that was dictating the tone of the music and the dance pose of the dancers. The ability of the machine to successfully classify the drumbeats is a pointer to the fact that the Traditional African drums are close to a time they will be digitized and beaten by the machine. It shows that technology is indeed able to unravel the “talking skill” of these drums.

5.1.4 Traditional African Dance Modelling

Deep learning techniques (pose estimation) is an effective way for generating dance pose stick model from a dance dataset. The pose estimation technique identified the important body joints of the human body; in this case seventeen (17) major joints and use the information to generate pose stick models. This technique is an effective way for digitizing dance poses. Some studies have explored the use of GANs for dance generation. Chan, Ginosar, Zhou, and Efros (2019) used a motion transfer method that could get poses from a dancer and transfer it to target in real time. This model was very successful, except that the dance was performed in real time. In this case the action was not performed to be digitized and stored in an archive. This model focused on motion transfer. Some other studies also targeted dance digitization, but the actions were recorded using motion sensors of various kinds. The use of motion sensors is common in motion studies, but it has the limitations of not being applicable where travel to get experts is limited. Another downside to the use of motion sensors is that the data collected is limited to few dancers. It is best when the dance documentation is focused on the dance styles of a particular performer and not on a dance type.

The pose sticks generated for the Sinte dance poses have been documented by the position of each of the joints per frame. The pose sticks can also be stringed together in a lightweight Graphics Interchanged Format (gif) that can be placed in documents various platforms. These pose sticks are available to programmers and graphics artist for use in different ways. The use of deep learning techniques has made it easy to record poses without having performers wearing sensors. This method expands the use of prerecorded videos for pose related studies.

5.1.5 Deep Framework for Intangible Heritage Framework

The proposed framework – Deep Culture Preservation Framework was formed based on the processes followed during this study. The field of artificial intelligence has not been explored much in the field of cultural preservation (Fiorucci, Khoroshiltseva, Pontil, Traviglia, Bue, and James, 2020). Most of the publicly available datasets are focused on photography, paintings, or music (Fiorucci et al., 2020). There seems to be less focus on the preservation of performing arts which is an important piece in the lives of people. From the experiences gathered along the course of this work a framework that can be replicated for the generation of data and further studies on the data using deep learning techniques is being documented. The Deep Culture Preservation Framework is an efficient framework for preserving intangible cultural heritages. It provided a model that can be adapted for any other deep study research on intangible cultural heritage, be it video or text dataset.

5.1.6 Contribution to Knowledge

The availability of this dataset and different explorations using deep learning techniques have implications for museums, culture enthusiasts, researchers, animators, and cultural preservation experts.

Implications for Museums

Museums are saddled with the responsibility of collecting, interpreting, displaying, and preserving cultural, scientific or artistic objects that are significant for the public. Thousands of people visit different museums to learn or experience new information. It is believed that this dataset will be an important collection for dance museums, art museums and even general museums. The videos are documentation of these dance performances by natives. Some of the videos provided information about the different dance poses. These videos can be retrieved for cultural education. These materials are not readily available for professionals in most places, and where they are, they are not indigenously performed. This study has solved the problem of availability and authenticity in its own little way.

The study also provided a framework for performance art digitization. Because of the advancement in technology, performing art preservation can now be taken farther than just recording videos and storing them up on shelves. The arts can be developed into phone applications. The arts can be stored in the cloud and can be manipulated to be made available to people beyond the four walls of the museum using readily available tools such as phones.

Implications for cultural enthusiasts

People who like to learn and experience other cultures usually have limited fulfilment until they travel to various places. Although tourists experience is intense when they actually travel, but at this stage technology can bring one as close as it can be. A Google search of ‘African Dances’ only returned few pages with some lists, images and a few links to some YouTube videos. The same video is referenced on most of the websites. A cause for this is probably that most indigenous people do not think much about the importance of documenting the culture (a major influence of westernization).

It is however a misconception to think that there are not cultural enthusiasts who wish to have the knowledge of this culture. In my opinion, there are many people who wish to know about other cultures without having to necessarily travel. As a matter of fact, the access to such information usually promotes travel. This study has contributed in its way to the information that is now available. It is hoped that within a short time from now there will be a website that will improve the visibility of this documentation. It is important that people find quality and accurate information about a culture. And, for tourism professionals, it is believed that making information about other cultures available will promote their business.

Another contribution of this study that has implications for cultural enthusiasts is that they can now have information on their phone when they travel. The dance classification model developed can be developed into a phone application. When an enthusiast has this application on their phones and have a dance performed when they travel, they can take a snap shot of the dancer and have the phone application provide information about what dance it is, the history of the dance, meaning of important poses and so many more details on the device as they move about. This eliminated the possibility of misinformation or fear when interacting with a foreign culture.

Implications for researchers

Two major categories of researchers will benefit from the results of this study – researchers in technology and researchers in the arts field. Although it is a great thing to develop new algorithms, it is also important to apply developed technology and examine its impact in other fields. It has been established in the recent time that interdisciplinary research has more impact in the society in which we live. Wen, Wang, Kozak, Liu, and Hou (2020) discussed the importance of carrying out interdisciplinary research. A major advantage is having real life applications for the theories that are developed.

For researchers in the field of artificial intelligence, this study provided a dataset that can be used for dance study or Traditional African dance study. Up until this study, there is no known publicly available dataset for Traditional African dance. Researchers who are interested in action recognition, dance recognition, video classification, object recognition and different other aspects of deep learning and machine learning have an available dataset to work with now. For researchers who are interested in robotic design, the pose estimation data used for the pose stick generation can be used for different robotic design. This data set can be used for different robotic design studies.

For researchers in the arts or performing arts, this study provides information about how deep learning can be applied in the performing arts. The developed framework provides a blueprint for digital preservation of intangible cultures using deep learning. The video datas so provides available information that can be used for a study on Traditional African dances or related studies. Researchers involved in digital preservation or dance digitization also have available information about how modern technology can be applied. Up until the time of this study, dance has been digitized using different notations such as labanotation or green notation. While these notations are informative, they are hard to understand by people who are not experts in deciphering the symbols. Another limitation with this method is it does not provide an exclusive visual information about the dance. It is more applicable to choreographies. So, while the dance is digitized, it is of very little use to people who are not experts in reading the symbols and patterns; thus, it is not available. This study provided pose stick images for each pose, and information about joint orientations which have more applications.

Implications for Animators

Animators are skilled in the manipulation of 2D and 3D motion data. The use of key framing in the animation of human motion is very popular in the computer animation field. This technique uses joint position information to construct a frame by frame sequence of action. The pose estimation data for Sinte dance can be applied with this technique to animate poses from the Sinte dance or the complete dance. This dataset provided a dancer's information in the digitized form. Several tools are now leveraging on artificial intelligence and deep learning techniques to achieve what took a lot of effort in the past in a short time and at better quality now. Adobe Character Animator is one of those tools. Character Animator can be used to animate a cartoon character with facial expression without a single line of code. A user can just use the application on their personal computer to capture facial expressions such as lip movement using the computer's camera and microphone and within a short time the expression is impressed on the cartoon.

The body joints motion of the dance can be used by experienced animators for various motion animation or generation.

Implications for Cultural Preservation Experts

Every aspect of this study has implications for experts in the field of cultural preservation. Intangible culture has been known to benefit less from the application of technology to cultural preservation. A major threat to all intangible cultural heritage was addressed by UNESCO to be weakened practice and transmission. Thus, every effort to strengthen practice and ensure transmission from one generation to another is a win for intangible cultural preservation.

A major contribution to strengthen practice is creating awareness about the culture. Culture globalization, another threat on its own, has reduced the awareness about what use to be popular. So even when people are aware about the practice, they become passive and soon lose the consciousness that the cultural practice exists. When made public, video collections can sensitize people to the existence of this culture.

Another important piece to this is that the information is made available for cultural learning. Since this is not a leisure study, a lot of important information was encountered during the research. Details about hand poses, which were not usually considered or understood by people who are not experts, were uncovered. History of the different dances studied were unraveled. Transformations during the dances were studied. This information is not only providing education for people from other cultures, but even natives. The study revealed several things that are known only to experts in dance performance.

Application of deep learning to Traditional African dance preservation is another contribution to this field. Up until the time of this study, no study was publicly available in published or unpublished form that focused on applying deep learning techniques applied to Traditional African dance preservation. There are a few studies carried out on Indian Classical dance but none in Traditional African dance. While some knowledge could be transferred, there are several differences in the performance of dances of Indian heritage and those of Africa. It is also important to state that until this study was carried out no known study used pose estimation to generate pose stick models from dance videos. This is a major contribution and application of an existing technique in an entirely different field of study. The poses generated are stored up in dictionaries that can be called upon and manipulated. It is hoped that similar techniques will be appropriately applied to the preservation of intangible cultures.

5.2 Future Work

The application of deep learning techniques to intangible cultural heritage is just beginning. Thus, there are many more areas that could be explored. There are several questions raised during this study that remain unanswered. Some of these aspects are in the pipeline for the researcher:

I. Development of documentation that will provide detailed information about the content of each video. This documentation can be used by anybody who has access to the videos and wants to understand what poses are used in the video and the meaning of those poses. The documentation will also be extended to the individual drums in each of the dance types.

II. Design of a website for the cultural resources on the Traditional African Dances. Crowdsourcing data collection is a good way of generating a diversified dataset. It is also believed that it will be easy to get a large dataset that may scale up to a size close to that of ImageNet using this technique. There are different types of websites that can be developed. A website can be developed for crowdsourcing dance videos while another will target just making the information available.

III. Development of a mobile (iOS and android) and web-based application on which this model can be deployed to make it available for tourists and people who have an interest in knowing more about Traditional African dance. The phone application will be useful for tourists as well as people who are interested in cultural education.

IV. The dance model can be expanded in scope by re-training the algorithm to identify more traditional dance forms. The dataset will also be expanded to provide information about more Traditional African dances.

V. Another research area is into carrying out the classification using Recurrent Neural Networks (RNNs). RNNs are notable for studies on time series data. Videos are examples of time series dataset, thus building classification model using RNNs can be experimented. This will provide further information about which technique has the best result.

VI. Another area of research is generating dance poses from the pose stick estimates using Generative Adversarial Networks (GANs). The aim may be generating a novel dance from existing dance poses. GANs have been applied in this capacity for music generation.

VII. Effort can also be made to manipulate the dance forms for action transfer using GANs.

VIII. Another important area of study is to design a tool that can be used for automated clipping of traditional videos. This tool will be of tremendous use in collecting videos of Traditional African dances.

5.3 Limitations

This study has some limitations which are:

- All the images generated for the classification process belong to only three classes- Adowa, Bata and Swange. This decision is based on the number of video available per class. Much more videos were available for these three classes than others.
- Pose stick generation from pose estimation algorithms was limited to only one dance class. The decision to use the dance class was because the costume used in the video did not occlude dancer's lower joints which was a major limitation with the other dance classes.
- Pose stick generation was that of a single dancer.

5.4 Summary

Conclusions based on the study as well as recommendations for future researchers were detailed in this chapter. Helpful suggestions which will make related studies easier for researchers were provided. Suggestions were also provided for further improvements on this work.

REFERENCES

Agber, T.C. (2013). Yamuel Yashi Agbatar Gyur: the pioneer of Swange music. Retrieved on 04/02/2021 from Tim Cuttings Agber — LiveJournal

Agyekum, K. (2010). The sociolinguistics of thanking in Akan. *Nordic Journal of African Studies 16*(2), 77-79

Anku, W. (2009). Drumming among the Akan and Anlo Ewe of Ghana: An introduction. African Music: *Journal of the International Library of African Music. 8*(3), 38-64

Arthur, K. (2006). *Reminiscence of ManhyiaTete Adowa: An original composition in contemporary idiom based on Adowa music of the Akans of Ghana.* [Unpublished MPhil thesis in Music]. University of Cape Coast. Retrieved on 03/10/2020 from (42) (PDF) Reminiscience of Manhyia Tete Adowa an original composition in contemporary idiom based on adowa music of the Akans of Ghana | Kras Arthur - Academia.edu

Bakalos, N., Rallis, I., Doulamis, N.D., Doulamis, A., Protopapadakis, E., and Voulodimos, A. (2019). Choreographic Pose Identification using Convolutional Neural Networks. *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, 1-7.

Berndt, E. and Carlos, J. (2000) "Cultural heritage in the mature era of computer graphics," in *IEEE Computer Graphics and Applications*, vol. 20, no. 1, pp. 36-37, Jan.-Feb. 2000.

Bloice, M.D., Stocker, C. and Holzinger, A. (2017) *Augmentor: An Image Augmentation Library for Machine Learning*. Retrieved from https://arxiv.org/pdf/1708.04680.pdf

Castro, D., Hickson, S., Sangkloy, P., Mittal, B., Dai, S., Hays, J., and Essa, I. (2018*) Let's Dance: Learning From Online Dance Videos*. Retrieved from https://arxiv.org/pdf/1801.07388.pdf

Cao, Z.; Hidalgo, G.; Simon, T.; Wei, S. -E. and Sheikh, Y. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 43*(1), 172-186.

Chan, C., Ginosar, S., Zhou, T., & Efros, A. A. (2019). *Everybody Dance Now*. arXiv.org. https://arxiv.org/abs/1808.07371.

Chianese, A. and Piccialli, F. (2014). "Designing a Smart Museum: When Cultural Heritage Joins IoT," *2014 Eighth International Conference on Next Generation Mobile Apps, Services and Technologies*, Oxford, 2014, pp. 300-306.

Cheng, Y., Peng, D., Sun, S. (2006) Creating a Virtual Activity for the Intangible Culture Heritage. In: *16th International Conference on Artificial Reality and Telexistence–Workshops, ICAT 2006*, pp. 636–641 (2006).

Cherian, A., Fernando, B., Harandi, M., Gould, S. (2017). Generalized rank pooling for activity recognition, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* Retrieved from https://arxiv.org/abs/1704.02112.

Chunwei, T., Lunke, F., Wenxian Z., Yong X., Wangmeng Z., and Chia-Wen L. (2020). Deep Learning on Image Denoising: An Overview. Retrieved from https://arxiv.org/pdf/1912.13171v2.pdf

Cignoni, P, and Roberto, S (2008) Sampled 3D models for CH applications: A viable and enabling new medium or just a technological exercise? *Journal of Computing Cultural Heritage* 1(1), 1–23 (2008).

Costa Y.M., Soares de Oliveira L.E. and Silla, C. (2017) An Evaluation of Convolutional Neural Networks for Music Classification Using Spectrograms. *Applied Soft Computing 52*. Retrieved from https://www.researchgate.net/publication/311650413_An_Evaluation_of_Convolutional_Neural_Networks_for_Music_Classification_Using_Spectrograms/citations

Feichtenhofer,C., Pinz, A. and Zisserman, A. (2016) Convolutional Two-Stream Network Fusion for Video Action Recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016, pp. 1933-1941.

Fernando, B., Anderson, P., Hutter, M. and Gould, S. (2016). Discriminative Hierarchical Rank Pooling for Activity Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016, pp. 1924-1932.

Fiorucci, M., Khoroshiltseva, M., Pontil, M., Traviglia, A., Bue, A. D., & James, S. (2020, February 17). *Machine Learning for Cultural Heritage: A Survey*. Pattern Recognition Letters. https://www.sciencedirect.com/science/article/pii/S0167865520300532.

Fowler P. 2002. World Heritage Cultural Landscapes, 1992-2002: A Review and Prospect. *Cultural landscapes: the Challenges of Conservation*, Paris, UNESCO , pp. 16-32.

Froggatt, A. (2006). Implementing a Corporate GIS to protect England's Heritage, 2006 ESRI International User Conference, 2006, (http://www.esri.com).

Gams M., and Tanea I. (2018) VR Technologies in Cultural Heritage. VRTCH 2018. *Communications in Computer and Information Science*, vol 904. Springer, Cham. Retrieved on March 3, 2020 from https://link.springer.com/chapter/10.1007/978-3-030-05819-7_8.

Geri, G.T. (2012). A history of Tiv religious practices and changes. A focus on Imborvungu, poor and Ibiamegh. Makurdi:Aboki Publishers.

Geron, A. (2019) *Hands-on Machine Learning with Scikit-Learn, Keras, and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd edition*, O'Reilly, ISBN-13: 978- 1492032649

Girdhar, R.., Ramanan, D., Gupta, A., Sivic, J. and Russell, B. (2017) Actionvlad: Learning spatio-temporal aggregation for action classification, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

Green, D. (2012). African oral tradition literacy. Journal of Black Studies.

Gupta S (2020) Indian Dance Form Classification Using Monk AI. Retrieved form Indian Dance Form Classification Using Monk AI | by Shubham Gupta | Towards Data Science on 03/25/2021

Haydar, M., Roussel, D., Maïdi, M., Otmane, S., and Mallem, M. (2011). Virtual and augmented reality for cultural computing and heritage: a case study of virtual exploration of underwater archaeological sites (preprint). *Virtual Reality 15*, 311–327. https://doi.org/10.1007/s10055-010-0176-4

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778).

Ikeuchi, K., Oishi, T., Takamatsu, J., Sagawa, R., Nakazawa, A., Kurazume, R., Nishino, K., Kamakura, M., Okamoto, Y. (2007) The Great Buddha Project: Digitally Archiving, Restoring, and Analyzing Cultural Heritage Objects. *International Journal of Computer Vision* 75(1), 189–208 (2007).

Ji, X. and Liu, H. (2010). Advances in View-Invariant Human Motion Analysis: A Review. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40*(1), 13-24.

Kar, A., Rai, N., Sikka, K. and Sharma, G. (2017). Adascan: Adaptive scan pooling in deep convolutional neural networks for human action recognition in videos. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R. and Li, F. (2014). Large-scale Video Classification with Convolutional Neural Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* pp. 1725-1732.

Kaushik, V., Mukherjee, P. and Lall, B., (2018) Nrityantar: Pose Oblivious Indian Classical Dance Sequence Classification System. Retrieved from https://arxiv.org/pdf/1812.05231.pdf

Kölb, O., Boussahl, M., and Hostettler, H. (2003). Requirements in an Inventory on Cultural Heritage in Morocco and Reflections on the Presentation of the Information, *CIPA 2003 XIXth International Symposium*, 30 September – 04 October 2003, Antalya, Turkey

Konsa K. (2013). Heritage as a Socio-Cultural Construct: Problems of Definition, *Baltic Journal of Art History 6*, p. 125.

Krishnan, K., Prabhu, N. and Babu, R. V. (2016) ARRNET: Action recognition through recurrent neural networks. *2,016 International Conference on Signal Processing and Communications (SPCOM)*, Bangalore, 2016, pp. 1-5.

Krizhevsky, A., Sustskever, I. and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems 25 (NIPS 2012)*

Kumar, K. V. V., Kishore, P. V. V., Kumar, D. A. and Kumar, E. K. (2018) Indian Classical Dance Action Identification Using Adaboost Multiclass Classifier on Multifeature Fusion. *2018 Conference on Signal Processing and Communication Engineering Systems (SPACES)*, Vijayawada, 2018, pp. 167-170.

LeCun, Y., Bengio, Y. and Hinton, G. (2015) "Deep learning," *Nature*, vol. *521* (7553) pp. 436–444.

LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. Gradient-Based Learning Applied to Document Recognition, *Proceedings of the IEEE, 86*(11):2278-2324.

Lu, Z., Yu, Z., Ya-Li, P., Shi-Gang, L., Xiaojun, W., Gang, L., and Yuan, R. (2018). Fast Single Image Super resolution Via Dilated Residual Networks. IEEE Access. Retrieved from https://www.researchgate.net/publication/327613660_Fast_Single_Image_Super-Resolution_via_Dilated_Residual_Networks

Ma C.-Y., Chen M.-H., Kira Z., AlRegib G. (2017) Ts-LSTM and Temporal-inception: Exploiting Spatiotemporal Dynamics for Activity Recognition. Retrieved from https://arxiv.org/abs/1703.10667

Markoff, J (November, 2012) Seeking a Better Way to Find Web Images The new York times. Retrieved from https://www.nytimes.com/2012/11/20/science/for-web-images-creating-new-technology-to-seek-and-find.html?_r=0

Moeslund, T.B.; and Granum,E. (2001). A survey of computer vision based human motion capture. *Computer Vision and Image Understanding*, *81*(3), 231-268.

Moeslund, T.B.; Hilton, A., and Kruger,V. (2006). A survey of advances ¨ in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, *104*(2-3), 90-126.

Park, E., Han, X., Berg, T. L. and Berg, A. C. (2016) Combining multiple sources of knowledge in deep CNNs for action recognition. *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Placid, NY, 2016, pp. 1-8.

Pavlidis, G., Koutsoudis, A., Arnaoutoglou, F., Tsioukas, V., Chamzas, C. (2007) Methods for 3D digitization of cultural heritage. *Journal of Cultural Heritage* 8(1), 93–98

Petrescu, F (2007). THE USE OF GIS TECHNOLOGY IN CULTURAL HERITAGE. XXI *International CIPA Symposium*, 01-06 October 2007, Athens, Greece. Pienaar, S.W. and Malekian R. (2019). Human Activity Recognition Using LSTM-RNN Deep Neural Network Architecture. Retrieved from https://arxiv.org/abs/1905.00599

Philbrick, K.A., Weston, A.D., Akkus, Z. Kline, T.L., Korfiatis, P., Sakinis, T.; Kostandy, P., Boonrod, A., Zeinoddini, A., Takahashi, N., and Erickson, B.J.(2019) RIL-Contour: a Medical Imaging Dataset Annotation Tool for and with Deep Learning. *Journal of Digital Imaging* **32,** 571–581 (2019). https://doi.org/10.1007/s10278-019-00232-0.

Poppe,R. (2007). Vision-based human motion analysis: An overview. *Computer Vision and Image Understanding*, *108*(1-2), 4-18.

Pugnaloni, F., Issini, G., Minh, N.D. (2008) 3D City Model of the Ancient Hue, Vietnam; Reconstruction of the City Environment for the Cultural Heritage Identity Conservation. In: Wyeld, T.G., Kenderdine, S., Docherty, M. (eds.) *VSMM* 2007. LNCS, vol. 4820, pp. 13–23. Springer, Heidelberg

Rafi, A., Salleh, A., Paul, A., Noraisah, S., Jun, Y., Hanif, R., Mahadzir, M. (2010) Modeling Optimization for Real-time Virtual Heritage Visualization Content: A research on e-Warisan SENIBINA. *International Symposium in Information Technology (ITSim)*, pp. 1–7

Raimi, K. (2019) Illustrated: 10 CNN Architectures. https://towardsdatascience.com/illustrated-10-cnn-architectures-95d78ace614d

Rahaman, H. and Tan, B.K. (2010) Interpreting Digital Heritage. *15th International Conference on Computer-Aided Architectural Design Research in Asia CAADRIA* 2010, pp. 93–102

Ramos, M. (2000). The empire beats on: Oyo, Bata drums and hegemony in nineteenth-century Cuba.

Reitz, T., Haist, J., Wigg-Wolf, D. (2006). Graphical Annotation of Semantic Units in a 3D GIS for Cultural Heritage Research and Presentation. *The evolution of Information Communication Technology in Cultural Heritage*, October 30 – November 4, 2006, Nicosia, Cyprus.

Shorten, C., Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *Journal Big Data* **6,** 60 (2019). https://doi.org/10.1186/s40537-019-0197-0

Simonyan, K., Zisserman, A. (2014a). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.

Simonyan, K. and Zisserman, A. (2014b) Two-Stream Convolutional Networks for Action Recognition in Videos. *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, pp. 568-576

Simonyan, K. and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *ICLR*, 2015.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015) Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015, pp. 1-9.

Takase, Y., Yano, K., Nakaya, T., Isoda, Y., Kawasumi, T., Matsuoka, K., Tanaka, S., Kawahara, N., Inoue, M., Tsukamoto, A., Kirimura, T., Kawahara, D., Sho, N., Shimiya, K., Sone, A., and Shiroki, M.(2006). Visualization of Historical City Kyoto by Applying VR and Web3D-GIS Technologies, *The evolution of Information Communication Technology in Cultural Heritage*, October 30 – November 4, 2006, Nicosia, Cyprus.

Toshev, A. and Szegedy,C. (2014). DeepPose: Human Pose Estimation via Deep Neural Networks. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014, pp. 1653-1660, doi: 10.1109/CVPR.2014.214.

Tsevende,R.A., Agber,T.C., Iorngurum, D.S, and Ugbagir, N.N. (2013). Tiv Swange Music and Dance. Midan Press Abuja,Nigeria.

UNESCO definition of cultural heritage. http://www.unesco.org. Accessed March 3, 2020.

Valtolina, S., Franzoni, S., Mazzoleni, P., and Bertino, E. (2005). Dissemination of Cultural Heritage Content through Virtual Reality and Multimedia Techniques: A Case Study. *Multimedia Modelling Conference (MMM 2005)*, pp. 214–221

Voinea GD., Girbacia F., Postelnicu C.C., and Marto A. (2019) Exploring Cultural Heritage Using Augmented Reality Through Google's Project Tango and ARCore. In: Duguleană M., Carrozzino

Wang, F. (2009) "Is Culture Computable?" *IEEE Intelligent Systems*, vol. 24, no. 2, pp. 2-3.

Wang, L., Qiao, Y. and Tang, X. (2015) Action recognition with trajectory-pooled deep-convolutional descriptors. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* pp. 4305-4314.

Wen, J., Wang, W., Kozak, M., Liu, X. and Hou, H. (2020) Many brains are better than one: the importance of interdisciplinary studies on COVID-19 in and beyond tourism, *Tourism Recreation Research*, DOI: 10.1080/02508281.2020.1761120

Xinyu, D., Baoqing, G., Chuangming, S. (2007) Virtual Exploration Application of Cultural Heritage for Anyang Yinxu. *Second Workshop on Digital Media and its Application in Museum & Heritages*, pp. 237–241 (2007)

Zhang N., Qi, G., Kieu, D. and Hua, K.A. (2019) Differential Recurrent Neural Network and its Application for Human Activity. Retrieved from https://arxiv.org/abs/1905.04293

Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L., 2017. Beyond A Gaussian Denoiser: Residual Learning of Deep CNN For Image Denoising. *IEEE Transactions on Image Processing 26* (7), 3142–3155.

Zhao, R., Ali, H., and van der Smagt, P. (2018) Two-stream RNN/CNN for Action Recognition In 3D Videos. *International Conference on Intelligent Robots and Systems (IROS), IEEE.* Retrieved online from https://arxiv.org/abs/1703.09783

Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y; Zhu, H.; Xiong, H.; and He, Q. (2020). A Comprehensive Survey on Transfer Learning. Retrieved on 04/09/2021 from 1911.02685.pdf (arxiv.org)

4kVideoDowloader (2020). Open Media LLC (version 4.14.1) [Computer Software]. Online. www.4kdownload.com/products/product-videodownloader

Aslam, 2021 • YouTube by the Numbers (2021): Stats, Demographics & Fun Facts (omnicoreagency.com)

Tzutalin. LabelImg. Git code (2015). https://github.com/tzutalin/labelImg

APPENDIX

/***

* This code is an adaptation of the large image dataset feature extraction project as referenced below.

*Title: Feature Extraction and Reverse Image Search

*Author: Dombrowski M.P.

*Date:2019

*Code version:

*Availability: Feature extraction and reverse image search – Nextjournal

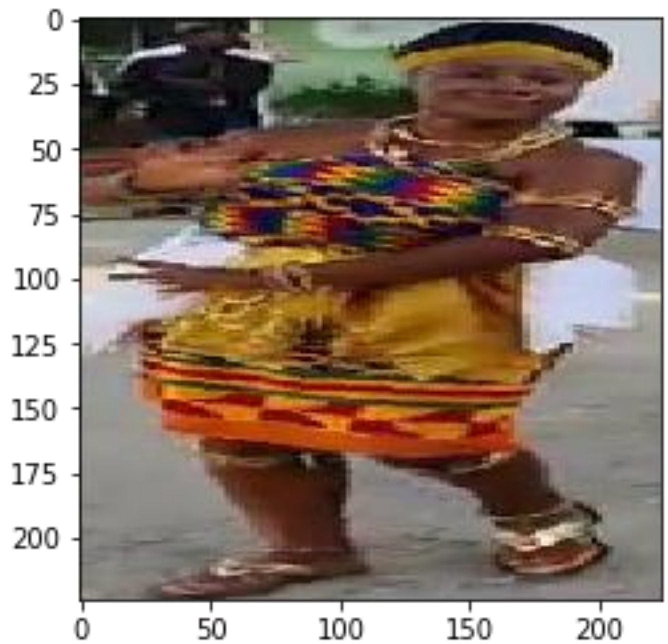
***/

```
import os
import keras
from keras.preprocessing import image
from keras.applications.imagenet_utils import decode_predictions,
preprocess_input
from keras.models import Model
In [2]:
model = keras.applications.VGG16(weights='imagenet', include_top=True)


In [3]:
model.summary()
In [5]:
%matplotlib inline
import numpy as np
import matplotlib.pyplot as plt

def load_image(path):
  img = image.load_img(path, target_size=model.input_shape[1:3])
  x = image.img_to_array(img)
  x = np.expand_dims(x, axis=0)
  x = preprocess_input(x)
  return img, x
In [6]:
img, x = load_image("D:\Dance Dataset 03222021A\Adowa\A24.jpg")
print("shape of x: ", x.shape)
print("data type: ", x.dtype)
```

```
plt.imshow(img)
```

Out[6]:

```
<matplotlib.image.AxesImage at 0x2106d89e6c8>
```



/** * Title:

In [7]:

```
# forward the image through the network
predictions = model.predict(x)

# print out the
for _, pred, prob in decode_predictions(predictions)[0]:
  print("predicted %s with probability %0.3f" % (pred, prob))
```

In [8]:

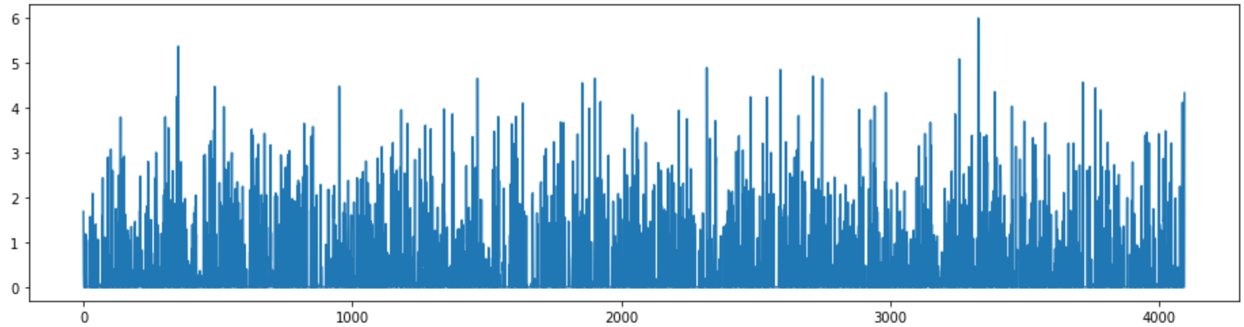
```
feat_extractor = Model(inputs=model.input,
outputs=model.get_layer("fc2").output)
feat_extractor.summary()
```

In [9]:

```
img, x = load_image("D:\Dance Dataset 03222021A\Adowa\A335.jpg")
feat = feat_extractor.predict(x)

plt.figure(figsize=(16,4))
plt.plot(feat[0])
```

Out[9]:

```
[<matplotlib.lines.Line2D at 0x2106d9b9c08>]
```



In [10]:

```
images_path = r"D:\Dance Dataset 03222021B"
image_extensions = ['.jpg', '.png', '.jpeg']  # case-insensitive (upper/lower
doesn't matter)
max_num_images = 10000

images = [os.path.join(dp, f) for dp, dn, filenames in os.walk(images_path)
for f in filenames if os.path.splitext(f)[1].lower() in image_extensions]
if max_num_images < len(images):
  images = [images[i] for i in sorted(random.sample(xrange(len(images)),
max_num_images))]

print("keeping %d images to analyze" % len(images))
```

```
keeping 7751 images to analyze
```

In [11]:

```
import time
tic = time.clock()


features = []
for i, image_path in enumerate(images):
  if i % 500 == 0:
    toc = time.clock()
    elap = toc-tic;
    print("analyzing image %d / %d. Time: %4.4f seconds." % (i,
len(images),elap))
    tic = time.clock()
  img, x = load_image(image_path);
  feat = feat_extractor.predict(x)[0]
  features.append(feat)
```

```
C:\Users\Adebunmi\anaconda3\lib\site-packages\ipykernel_launcher.py:2:
DeprecationWarning: time.clock has been deprecated in Python 3.3 and will be
removed from Python 3.8: use time.perf_counter or time.process_time instead
```

```
C:\Users\Adebunmi\anaconda3\lib\site-packages\ipykernel_launcher.py:8:
DeprecationWarning: time.clock has been deprecated in Python 3.3 and will be
removed from Python 3.8: use time.perf_counter or time.process_time instead

C:\Users\Adebunmi\anaconda3\lib\site-packages\ipykernel_launcher.py:11:
DeprecationWarning: time.clock has been deprecated in Python 3.3 and will be
removed from Python 3.8: use time.perf_counter or time.process_time instead
 # This is added back by InteractiveShellApp.init_path()
analyzing image 0 / 7751. Time: 0.0005 seconds.
analyzing image 500 / 7751. Time: 236.8060 seconds.
analyzing image 1000 / 7751. Time: 277.6630 seconds.
analyzing image 1500 / 7751. Time: 166.3325 seconds.
analyzing image 2000 / 7751. Time: 164.7794 seconds.
analyzing image 2500 / 7751. Time: 168.3971 seconds.
analyzing image 3000 / 7751. Time: 195.4713 seconds.
analyzing image 3500 / 7751. Time: 211.1208 seconds.
analyzing image 4000 / 7751. Time: 218.6097 seconds.
analyzing image 4500 / 7751. Time: 208.5042 seconds.
analyzing image 5000 / 7751. Time: 167.4900 seconds.
analyzing image 5500 / 7751. Time: 172.8269 seconds.
analyzing image 6000 / 7751. Time: 167.2116 seconds.
analyzing image 6500 / 7751. Time: 171.0625 seconds.
analyzing image 7000 / 7751. Time: 177.4620 seconds.
analyzing image 7500 / 7751. Time: 170.6728 seconds.
```

In [12]:

```
from sklearn.decomposition import PCA

features = np.array(features)
pca = PCA(n_components=300)
pca.fit(features)
```

Out[12]:

```
PCA(copy=True, iterated_power='auto', n_components=300, random_state=None,
  svd_solver='auto', tol=0.0, whiten=False)
```

In [13]:

```
pca_features = pca.transform(features)
```

In [14]:

```
import random

# grab a random query image
query_image_idx = int(len(images) * random.random())

# let's display the image
img = image.load_img(images[query_image_idx])
```

```
plt.imshow(img)
```

Out[14]:

```
<matplotlib.image.AxesImage at 0x21002014088>
```

In [15]:

```
from scipy.spatial import distance

similar_idx = [ distance.cosine(pca_features[query_image_idx], feat) for feat
in pca_features ]
```

In [16]:

```
idx_closest = sorted(range(len(similar_idx)), key=lambda k:
similar_idx[k])[1:6]
```

In [24]:

```
# load all the similarity results as thumbnails of height 100
thumbs = []
for idx in idx_closest:
  img = image.load_img(images[idx])
  img = img.resize((int(img.width * 100 / img.height), 100))
  thumbs.append(img)

# concatenate the images into a single image
concat_image = np.concatenate([np.asarray(t) for t in thumbs], axis=1)

# show the image
plt.figure(figsize = (16,12))
plt.imshow(concat_image)
plt.savefig('pca')
```

In [25]:

```
def get_closest_images(query_image_idx, num_results=5):
  distances = [ distance.cosine(pca_features[query_image_idx], feat) for feat
in pca_features ]
  idx_closest = sorted(range(len(distances)), key=lambda k:
distances[k])[1:num_results+1]
  return idx_closest

def get_concatenated_images(indexes, thumb_height):
  thumbs = []
  for idx in indexes:
    img = image.load_img(images[idx])
    img = img.resize((int(img.width * thumb_height / img.height),
thumb_height))
    thumbs.append(img)
  concat_image = np.concatenate([np.asarray(t) for t in thumbs], axis=1)
  return concat_image
```

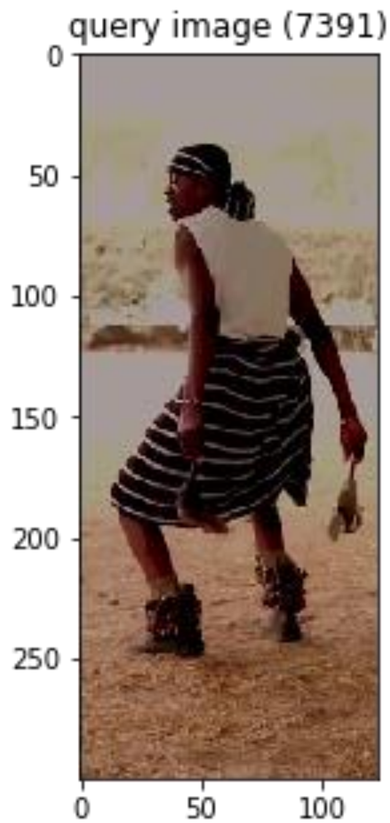
In [26]:

```
# do a query on a random image
query_image_idx = int(len(images) * random.random())
idx_closest = get_closest_images(query_image_idx)
query_image = get_concatenated_images([query_image_idx], 300)
results_image = get_concatenated_images(idx_closest, 200)
```

```
# display the query image
plt.figure(figsize = (5,5))
plt.imshow(query_image)
plt.title("query image (%d)" % query_image_idx)
```

Out[26]:

```
Text(0.5, 1.0, 'query image (7391)')
```



In [27]:

```
# display the resulting images
plt.figure(figsize = (16,12))
plt.imshow(results_image)
plt.title("result images")
plt.savefig('pca2')
```

In [28]:

```
# do a query on a random image
query_image_idx = int(len(images) * random.random())
idx_closest = get_closest_images(query_image_idx)
query_image = get_concatenated_images([query_image_idx], 300)
results_image = get_concatenated_images(idx_closest, 200)

# display the query image
plt.figure(figsize = (5,5))
plt.imshow(query_image)
plt.title("query image (%d)" % query_image_idx)
```

Out[28]:

```
Text(0.5, 1.0, 'query image (6462)')
```

In [29]:

```
# display the resulting images
plt.figure(figsize = (16,12))
plt.imshow(results_image)
```

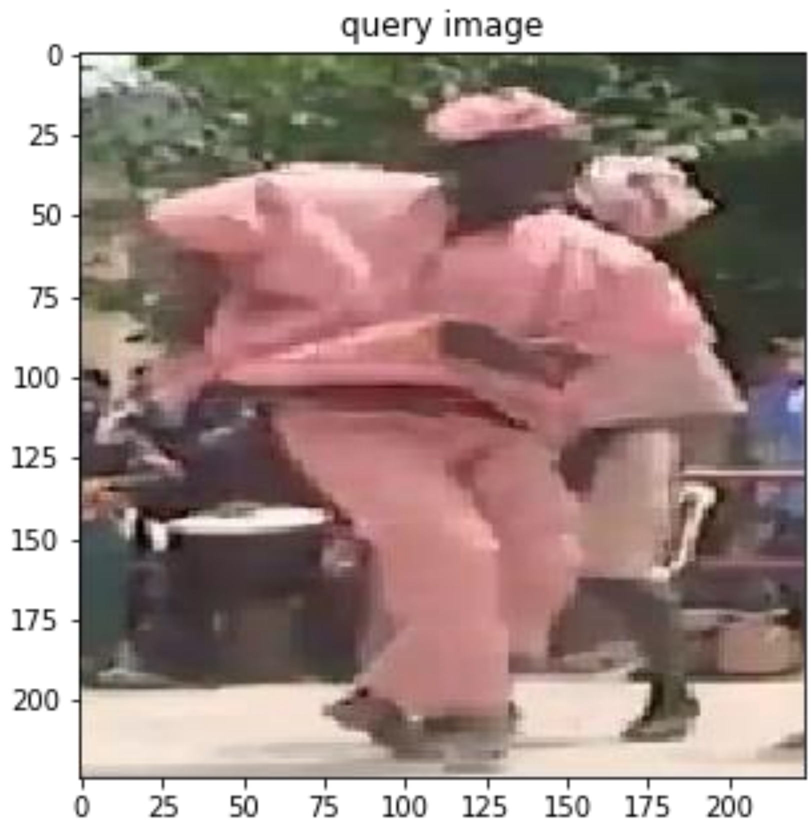
```
plt.title("result images")
plt.savefig('pca3')
```

In [30]:

```
#from google.colab import drive
#drive.mount('/content/gdrive', force_remount=True)

# load image and extract features
new_image, x = load_image(r"D:\Dance Dataset 03222021A\Bata\B23.jpg")
new_features = feat_extractor.predict(x)

# project it into pca space
new_pca_features = pca.transform(new_features)[0]

# calculate its distance to all the other images pca feature vectors
distances = [ distance.cosine(new_pca_features, feat) for feat in
pca_features ]
idx_closest = sorted(range(len(distances)), key=lambda k: distances[k])[0:5]
# grab first 5
results_image = get_concatenated_images(idx_closest, 200)

# display the results
plt.figure(figsize = (5,5))
plt.title("query image")
plt.imshow(new_image)
```

Out[30]:

```
<matplotlib.image.AxesImage at 0x2100bcfe948>
```



In [31]:

```
# display the resulting images
plt.figure(figsize = (16,12))
plt.title("result images")
plt.imshow(results_image)
plt.savefig('pca4')
```

In [32]:

```
import pickle

pickle.dump([images, pca_features, pca],
open(r"C:\Users\Adebunmi\features_dancedata101.p", 'wb'))
```

B

/***

*This code is an adaptation of a CNN model as referenced below.

*Title: Loading in your own data- Deep Learning basicsd with Python, Tensorflow and Keras p.2

*Author: Sendtex

*Date:2018

*Code version:

*Availability: [Python Programming Tutorials]()

***/

```
import math  # for mathematical operations
import matplotlib.pyplot as plt  # for plotting the images
%matplotlib inline
import pandas as pd
import os
import cv2
from keras.preprocessing import image  # for preprocessing the images
import numpy as np  # for mathematical operations
from keras.utils import np_utils
from skimage.transform import resize  # for resizing images
import tensorflow.keras as keras
from tensorflow.keras.models import Sequential
from keras.layers import Activation
from tensorflow.keras.layers import Conv2D, Flatten, MaxPooling2D, Dropout,
Dense, BatchNormalization
```

In [ ]:

```
DATADIR = r"D:\Dance Dataset 03222021A"
CATEGORIES = ["Adowa", "Bata", "Swange"]
```

In [ ]:

```
for category in CATEGORIES:
  path = os.path.join(DATADIR,category)
  for img in os.listdir(path):
    img_array = cv2.imread(os.path.join(path,img))
    plt.imshow(img_array)
    plt.show()
    break
  break
```

In [ ]:
```
print(img_array)
```

In [ ]:
```
print(img_array.shape)
```

In [ ]:
```
IMG_SIZE = 224

new_array = cv2.resize(img_array, (IMG_SIZE, IMG_SIZE))
plt.imshow(new_array)
plt.show()
```

In [ ]:
```
training_data = []

def create_training_data():
  for category in CATEGORIES:
    path = os.path.join(DATADIR,category)
    class_num = CATEGORIES.index(category)
    for img in os.listdir(path):
      try:
        img_array = cv2.imread(os.path.join(path,img))
        new_array = cv2.resize(img_array, (IMG_SIZE, IMG_SIZE))
        training_data.append([new_array,class_num])
      except Exception as e:
        pass
create_training_data()
```

In [ ]:
```
print(len(training_data))
```

In [ ]:
```
import random
random.shuffle(training_data)
```

In [ ]:
```
for sample in training_data[:10]:
  print(sample[1])
```

In [ ]:
```
X = []
y = []
```

In [ ]:
```
for features,label in training_data:
  X.append(features)
  y.append(label)
```

In [ ]:
```
import pickle

pickle_out = open("X.pickle","wb")
pickle.dump(X, pickle_out)
pickle_out.close()

pickle_out = open("y.pickle","wb")
pickle.dump(y, pickle_out)
pickle_out.close()
```

In [ ]:
```
pickle_in = open("X.pickle","rb")
X = pickle.load(pickle_in)
```

In [ ]:
```
X[1]
```

In [ ]:
```
X = pickle.load(open("X.pickle", "rb"))
y = pickle.load(open("y.pickle", "rb"))
```

In [ ]:
```
X = np.array(X).reshape(-1, IMG_SIZE, IMG_SIZE, 3)
```

X = np.asarray(X)/np.asarray(255.0) print(X.shape) y = np.asarray(y)/np.asarray(255.0) print(y.shape)
from sklearn.model_selection import train_test_split X_train, X_valid, y_train, y_valid = train_test_split(X, y, test_size=0.3, random_state=42)

In [ ]:
```
from keras.applications import ResNet50
```

base_model = ResNet50(weights='imagenet', include_top=False, input_shape=(150, 150, 3))

X_train = base_model.predict(X)

In [ ]:
```
model = ResNet50(weights='imagenet', include_top=True, input_shape=(224, 224,
3))
```

In [ ]:
```
#model.compile(optimizers.rmsprop(lr=0.0001, decay=1e-
6),loss="binary_crossentropy",metrics=["accuracy"])
model.compile(loss="sparse_categorical_crossentropy",
       optimizer="sgd",
       metrics=["accuracy"])
```

In [ ]:
```
history = model.fit(X,y,batch_size=32,epochs=50,validation_split=0.2)
```

In [ ]:
```
# Final evaluation of the model
```

```
scores = model.evaluate(X, y, verbose=0)
print("Accuracy: %.2f%%" % (scores[1]*100))
```

```
In [ ]:
# Save model weights
model.save_weights('weights_epoch_30.h5')
```

```
In [ ]:
# list all data in history
print(history.history.keys())
# summarize history for accuracy
plt.plot(history.history['accuracy'])
plt.plot(history.history['val_accuracy'])
plt.title('model accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='upper left')
plt.show()
# summarize history for loss
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='upper left')
plt.show()
```

C

/**

*This code is an adaptation of the running pose estimation project as referenced below.

*Title: Running Pose Estimate

*Author: Mader K.S.

*Date:2019

*Code version:

*Availability: Running Pose Estimate | Kaggle

**/

The kernel shows how to use the tf_pose_estimation package in Python on a series of running videos.

Libraries we need

Install tf_pose and pycocotools

In [1]:

In [2]:

```
!pip install -qq pycocotools
```

In [4]:

```
%load_ext autoreload
%autoreload 2
import seaborn as sns
import matplotlib.pyplot as plt
plt.rcParams["figure.figsize"] = (8, 8)
plt.rcParams["figure.dpi"] = 125
plt.rcParams["font.size"] = 14
plt.rcParams['font.family'] = ['sans-serif']
plt.rcParams['font.sans-serif'] = ['DejaVu Sans']
plt.style.use('ggplot')
sns.set_style("whitegrid", {'axes.grid': False})
```

```
The autoreload extension is already loaded. To reload it, use:
  %reload_ext autoreload
```

In [5]:

```
%matplotlib inline
import tf_pose
import cv2
from glob import glob
from tqdm import tqdm_notebook
from PIL import Image
import numpy as np
import os
def video_gen(in_path):
    c_cap = cv2.VideoCapture(in_path)
    while c_cap.isOpened():
        ret, frame = c_cap.read()
        if not ret:
            break
        yield c_cap.get(cv2.CAP_PROP_POS_MSEC), frame[:, :, ::-1]
    c_cap.release()
```

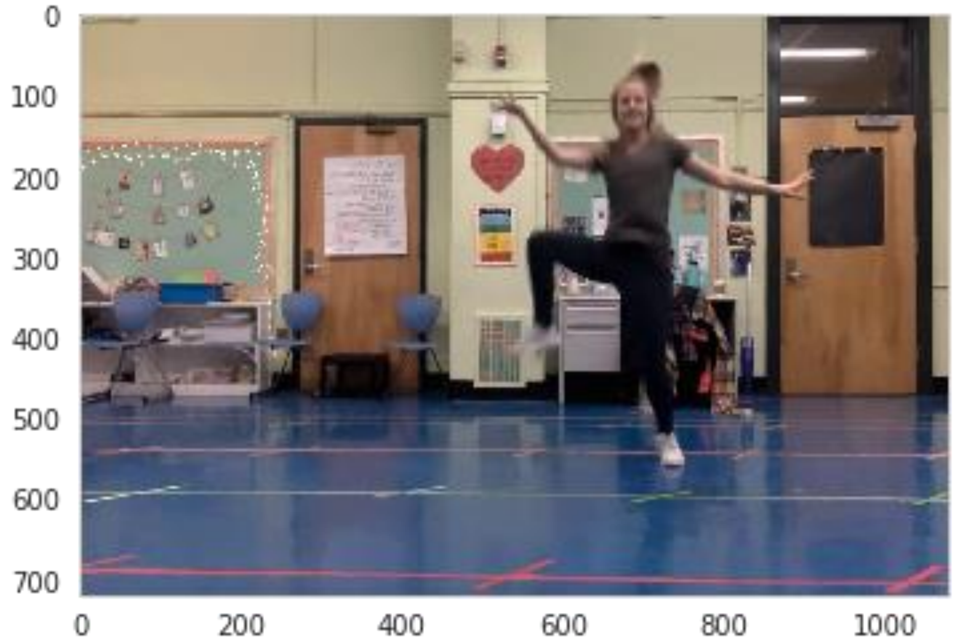
In [10]:

```
video_paths = glob('../input/sinte-dance-elite/Sinte  Dance Elite.mp4')
#video_paths = glob('../input/damian/*.mp4')
c_video = video_gen(video_paths[0])
for _ in range(300):
    c_ts, c_frame = next(c_video)
```

```
plt.imshow(c_frame)
```

Out[10]:

```
<matplotlib.image.AxesImage at 0x7f0de7899390>
```



In [11]:

```
from tf_pose.estimator import TfPoseEstimator
from tf_pose.networks import get_graph_path, model_wh
tfpe = tf_pose.get_estimator()
```

```
[2021-02-11 21:56:18,120] [TfPoseEstimator] [INFO] loading graph from
/opt/conda/lib/python3.6/site-
packages/tf_pose_data/graph/cmu/graph_opt.pb(default size=432x368)
2021-02-11 21:56:18,120 INFO loading graph from
/opt/conda/lib/python3.6/site-
packages/tf_pose_data/graph/cmu/graph_opt.pb(default size=432x368)
```

In [12]:

```
humans = tfpe.inference(npimg=c_frame, upsample_size=4.0)
print(humans)
```

```
[BodyPart:0-(0.63, 0.16) score=0.91 BodyPart:1-(0.64, 0.23) score=0.92
BodyPart:2-(0.60, 0.23) score=0.83 BodyPart:3-(0.54, 0.24) score=0.76
BodyPart:4-(0.50, 0.18) score=0.64 BodyPart:5-(0.68, 0.23) score=0.91
BodyPart:6-(0.74, 0.28) score=0.76 BodyPart:7-(0.81, 0.29) score=0.61
BodyPart:8-(0.60, 0.41) score=0.72 BodyPart:9-(0.53, 0.39) score=0.73
BodyPart:10-(0.53, 0.53) score=0.75 BodyPart:11-(0.65, 0.42) score=0.67
BodyPart:12-(0.66, 0.59) score=0.72 BodyPart:13-(0.67, 0.72) score=0.82
BodyPart:14-(0.62, 0.15) score=0.91 BodyPart:15-(0.64, 0.15) score=0.89
BodyPart:16-(0.61, 0.16) score=0.56 BodyPart:17-(0.66, 0.15) score=0.72]
```

In [13]:

```
new_image = TfPoseEstimator.draw_humans(c_frame[:, :, ::-1], humans,
imgcopy=False)
fig, ax1 = plt.subplots(1, 1, figsize=(10, 10))
ax1.imshow(new_image[:, :, ::-1])
```

Out[13]:

```
<matplotlib.image.AxesImage at 0x7f0dd6f860f0>
```

In [14]:

```
body_to_dict = lambda c_fig: {'bp_{}_{}'.format(k, vec_name): vec_val
                              for k, part_vec in c_fig.body_parts.items()
                              for vec_name, vec_val in zip(['x', 'y',
'score'],
                                                           (part_vec.x, 1-
part_vec.y, part_vec.score))}
c_fig = humans[0]
body_to_dict(c_fig)
```

Out[14]:

```
{'bp_0_x': 0.6296296296296297,
 'bp_0_y': 0.842391304347826,
 'bp_0_score': 0.9075720906257629,
 'bp_1_x': 0.6388888888888888,
 'bp_1_y': 0.7717391304347826,
 'bp_1_score': 0.9162707328796387,
```

'bp_2_x': 0.6018518518518519,
'bp_2_y': 0.7717391304347826,
'bp_2_score': 0.834660530090332,
'bp_3_x': 0.5416666666666666,
'bp_3_y': 0.7608695652173914,
'bp_3_score': 0.7607961893081665,
'bp_4_x': 0.5046296296296297,
'bp_4_y': 0.8206521739130435,
'bp_4_score': 0.6363097429275513,
'bp_5_x': 0.6759259259259259,
'bp_5_y': 0.7717391304347826,
'bp_5_score': 0.9108472466468811,
'bp_6_x': 0.7361111111111112,
'bp_6_y': 0.7228260869565217,
'bp_6_score': 0.7615177035331726,
'bp_7_x': 0.8101851851851852,
'bp_7_y': 0.7065217391304348,
'bp_7_score': 0.6132577657699585,
'bp_8_x': 0.6018518518518519,
'bp_8_y': 0.5923913043478262,
'bp_8_score': 0.7169932126998901,
'bp_9_x': 0.5277777777777778,
'bp_9_y': 0.6141304347826086,
'bp_9_score': 0.7349910736083984,
'bp_10_x': 0.5324074074074074,
'bp_10_y': 0.47282608695652173,
'bp_10_score': 0.7490565180778503,
'bp_11_x': 0.6481481481481481,
'bp_11_y': 0.5760869565217391,
'bp_11_score': 0.6652555465698242,
'bp_12_x': 0.6574074074074074,
'bp_12_y': 0.4130434782608695,
'bp_12_score': 0.7154346704483032,
'bp_13_x': 0.6712962962962963,
'bp_13_y': 0.27717391304347827,
'bp_13_score': 0.8217236995697021,
'bp_14_x': 0.6203703703703703,
'bp_14_y': 0.8532608695652174,
'bp_14_score': 0.9057081937789917,
'bp_15_x': 0.6388888888888888,
'bp_15_y': 0.8532608695652174,
'bp_15_score': 0.8873612880706787,
'bp_16_x': 0.6111111111111112,
'bp_16_y': 0.842391304347826,

```
 'bp_16_score': 0.5647059082984924,
 'bp_17_x': 0.6574074074074074,
 'bp_17_y': 0.8478260869565217,
 'bp_17_score': 0.7238561511039734}
```

In [15]:

```
MAX_FRAMES = 3200
body_pose_list = []
for vid_path in tqdm_notebook(video_paths, desc='Files'):
    c_video = video_gen(vid_path)
    c_ts, c_frame = next(c_video)
    out_path = '{}_out.avi'.format(os.path.split(vid_path)[1])
    out = cv2.VideoWriter(out_path,
                          cv2.VideoWriter_fourcc('M','J','P','G'),
                          10,
                          (c_frame.shape[1], c_frame.shape[0]))
    for (c_ts, c_frame), _ in zip(c_video,
                                  tqdm_notebook(range(MAX_FRAMES),
desc='Frames')):
        bgr_frame = c_frame[:,:,::-1]
        humans = tfpe.inference(npimg=bgr_frame, upsample_size=4.0)
        for c_body in humans:
            body_pose_list += [dict(video=out_path, time=c_ts,
**body_to_dict(c_body))]
        new_image = TfPoseEstimator.draw_humans(bgr_frame, humans,
imgcopy=False)
        out.write(new_image)
    out.release()
```

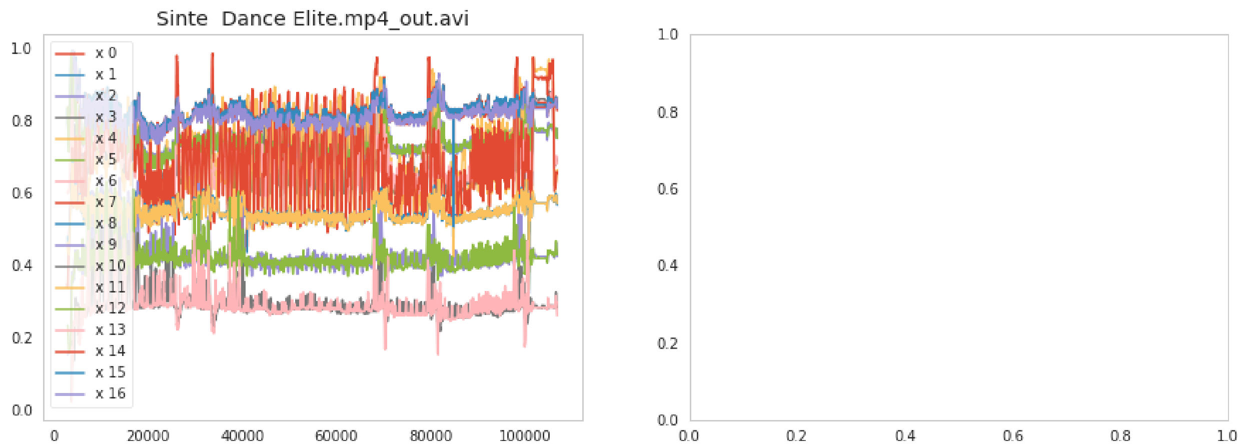
In [16]:

```
import pandas as pd
body_pose_df = pd.DataFrame(body_pose_list)
body_pose_df.describe()
```

Out[16]:

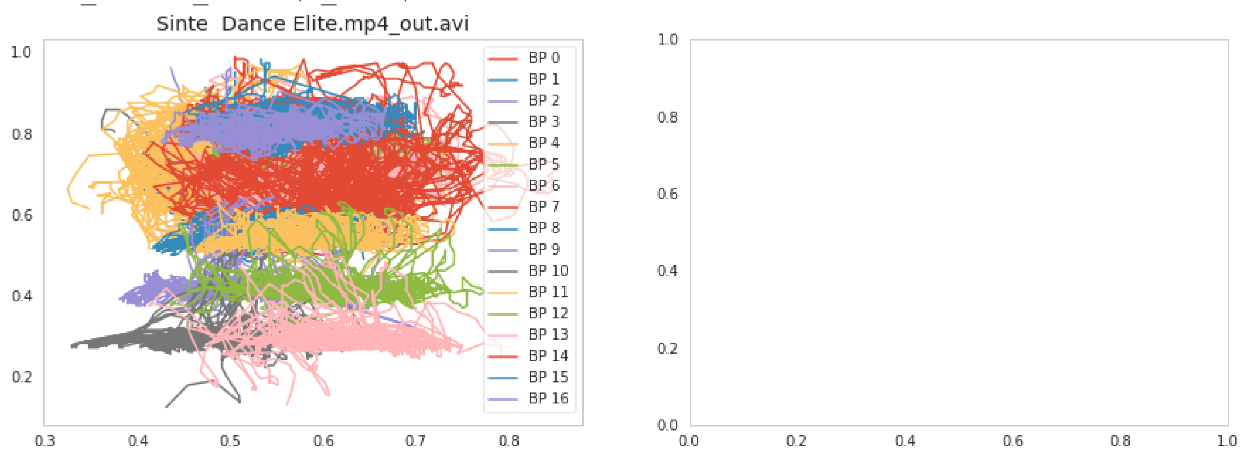
In [17]:

```
fig, m_axs = plt.subplots(1, 2, figsize=(15, 5))
for c_ax, (c_name, c_rows) in zip(m_axs, body_pose_df.groupby('video')):
    for i in range(17):
        c_ax.plot(c_rows['time'], c_rows['bp_{}_y'.format(i)], label='x
{}'.format(i))
    c_ax.legend()
    c_ax.set_title(c_name)
```



In [18]:

```
fig, m_axs = plt.subplots(1, 2, figsize=(15, 5))
for c_ax, (c_name, n_rows) in zip(m_axs, body_pose_df.groupby('video')):
    for i in range(17):
        c_rows = n_rows.query('bp_{}_score>0.6'.format(i)) # only keep confident results
        c_ax.plot(c_rows['bp_{}_x'.format(i)], c_rows['bp_{}_y'.format(i)], label='BP {}'.format(i))
    c_ax.legend()
    c_ax.set_title(c_name)
```



In [21]:

```
body_pose_df.to_csv('body_poseesti.csv', index=False)
```

In [ ]: