

Winter 12-13-2017

# Context-Sensitive Human Activity Classification in Video Utilizing Object Recognition and Motion Estimation

Abigail R. Jacoby

*University of New Mexico - Main Campus*

Follow this and additional works at: [https://digitalrepository.unm.edu/ece\\_etds](https://digitalrepository.unm.edu/ece_etds)



Part of the [Electrical and Computer Engineering Commons](#)

---

## Recommended Citation

Jacoby, Abigail R.. "Context-Sensitive Human Activity Classification in Video Utilizing Object Recognition and Motion Estimation." (2017). [https://digitalrepository.unm.edu/ece\\_etds/404](https://digitalrepository.unm.edu/ece_etds/404)

This Thesis is brought to you for free and open access by the Engineering ETDs at UNM Digital Repository. It has been accepted for inclusion in Electrical and Computer Engineering ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact [disc@unm.edu](mailto:disc@unm.edu).

**Abigail Jacoby**

*Candidate*

**Electrical and Computer Engineering**

*Department*

This thesis is approved, and it is acceptable in quality and form for publication:

*Approved by the Thesis Committee:*

**Prof. Marios Pattichis**

, Chairperson

**Prof. Manel Martinez-Ramon**

**Dr. Ramiro Jordan**

---

---

---

---

---

---

---

---

---

---

# Context-Sensitive Human Activity Classification in Video Utilizing Object Recognition and Motion Estimation

by

**Abigail Ruth Jacoby**

B.S., Computer Engineering, University of New Mexico, 2016

THESIS

Submitted in Partial Fulfillment of the  
Requirements for the Degree of

Master of Science  
Computer Engineering

The University of New Mexico

Albuquerque, New Mexico

May, 2018

# Dedication

*In memorium of my Aunt Linda, Uncles Gordon and Steve,  
and Grandmother Lois whom we've lost this semester,  
and of my Aunt Helen who passed 3 years recent.*

*To everyone who said I'll never make it...*

*"I've always been a fighter. If you tell me I can't,  
I'll die trying to prove you wrong."*

*– R. A. Salvatore*

# Acknowledgments

I thank my advisor and labmates for putting up with me this year, and my friends who have been indispensable to me while I was being an insufferable bore. Also to my cats, netflix, and actual gallons of coffee, I couldn't have done it without you.

A special thank you is extended to the National Science Foundation for funding the AOLME project<sup>1</sup> and for everything else they do to empower underrepresented groups in STEM.

---

<sup>1</sup>This material is based upon work supported by the National Science Foundation under Grant No. 1613637 and NSF AWD CNS-1422031.

# Context-Sensitive Human Activity Classification in Video Utilizing Object Recognition and Motion Estimation

by

**Abigail Ruth Jacoby**

B.S., Computer Engineering, University of New Mexico, 2016

M.S., Computer Engineering, University of New Mexico, 2017

## **Abstract**

Classification of human activity in raw video presents a challenging problem that remains unsolved, and is of great interest for large datasets. Though there have been several attempts at applying image processing techniques to video to recognize human activity in controlled video segments, few have attained a significant degree of success in raw videos.

Raw video classification exhibits significant challenges that can be addressed through the use of geometric information. Current techniques employ a combination of temporal information of the feature space or a combination of Convolutional and Recurrent Neural Networks (CNN and RNNs). CNNs are used for frame feature extraction and RNNs are then applied for motion vector extraction and classification. These techniques, which utilize information from the entirety of a frame, attempt to classify action based on all motion vectors and all objects found in the video. Such methods are cumbersome, often difficult to train, and do not generalize well beyond the dataset used.

This thesis explores the use of color based object detection in conjunction with contextualization of object interaction to isolate motion vectors specific to an activity sought within uncropped video. Feature extraction in this thesis differs significantly from other methods by using geometric relationships between objects to infer context. The approach avoids the need for video cropping or substantial preprocessing by significantly reducing the number of features analyzed in a single frame. The method was tested using 43 uncropped video clips with 620 video frames for writing, 1050 for typing, and 1755 frames for talking. Using simple KNN classification, the method gave accuracies of 72.6% for writing, 71% for typing and 84.6% for talking. Classification accuracy improved to 92.5% (writing), 82.5% (typing) and 99.7% (talking) with the use of a trained Deep Neural Network.

# Contents

<b>List of Figures</b>	<b>x</b>
<b>List of Tables</b>	<b>xiv</b>
<b>Glossary</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Overview . . . . .	1
1.2 Motivation . . . . .	2
1.3 Thesis Statement . . . . .	4
1.4 Contributions . . . . .	5
1.5 Summary . . . . .	5
<b>2 Background</b>	<b>7</b>
2.1 Classification of Human Activity in Video . . . . .	13
<b>3 Methodology</b>	<b>16</b>



## Contents

3.1	Overview of the Approach . . . . .	16
3.2	Candidate Region Selection Based on Color Models . . . . .	17
3.2.1	Color Models for Pencils, Table, Paper, Keyboard, and Faces	17
3.3	Context-based Processing and Feature Extraction . . . . .	21
3.3.1	KNN Classifier for Selecting Keyboard and Face Components	21
3.3.2	Writing . . . . .	22
3.3.3	Typing . . . . .	28
3.3.4	Talking . . . . .	32
3.3.5	Feature vectors . . . . .	36
3.4	Classification . . . . .	36
<b>4</b>	<b>Results</b>	<b>37</b>
4.1	Dataset . . . . .	37
4.2	Classification Model Results . . . . .	38
4.2.1	Results for Writing . . . . .	38
4.2.2	Results for Typing . . . . .	40
4.2.3	Results for Talking . . . . .	43
4.3	Failed Classification Examples . . . . .	45
4.4	Discussion . . . . .	52
<b>5</b>	<b>Conclusions and Future Work</b>	<b>56</b>

*Contents*

5.1	Conclusions . . . . .	56
5.2	Future Work . . . . .	58
<b>6</b>	<b>Appendices</b>	<b>60</b>
	<b>Appendices</b>	<b>60</b>
<b>A</b>	<b>Python Code for Color Detection</b>	<b>61</b>
A.1	Paper . . . . .	61
A.2	Pencil . . . . .	62
A.3	Keyboard . . . . .	63
A.4	Table . . . . .	64
A.5	Skin . . . . .	64

# List of Figures

1.1	Writing (a) v.s Not Writing (b), Typing (c) v.s. Not Typing (d), and Talking (e) v.s. Not Talking (f) . . . . .	3
1.2	Cropped video sample. . . . .	4
3.1	Overview flowchart. . . . .	16
3.2	Masked pencil color (right) v.s. original frame (left). . . . .	18
3.3	Subfigure a shows the collection of pencil variations we see, and b shows what we wish to achieve after masking and thresholding. . . .	18
3.4	A cropped writing video frame with potential pencil highlighted. . .	19
3.5	Flowchart displaying the process of extracting colors from a frame. .	20
3.6	A sample of faces that are used to train the face KNN model. . . . .	22
3.7	A collection of "Non-Faces" used in training the face KNN model. .	22
3.8	A collection of figures used in training the keyboard KNN, with keyboard samples in (3.8a),(3.8b), and (3.8c), and non-keyboards (monitors in in (3.8d), (3.8e) and mice in (3.8f), (3.8g), (3.8h)) . . . . .	24
3.9	A correctly identified keyboard using KNN . . . . .	25

*List of Figures*

3.10	A correctly identified face in KNN. . . . .	25
3.11	Paper detection by color, intermediate steps. . . . .	26
3.12	Pencil detection by color, intermediate steps. . . . .	26
3.13	Flowchart for confirmation of a pencil's presence prior to motion vector extraction. . . . .	27
3.14	Pseudocode for function <b>Context()</b> part 1, writing. . . . .	28
3.15	Images used as input for masking the background of the keyboard image . . . . .	29
3.16	Figures (a) and (b) display the two components required to determine if typing may be present. . . . .	29
3.17	A keyboard which has been properly found and classified for motion vector extraction. . . . .	30
3.18	A flowchart representing the steps to extracting the keyboard and discovery of potential typing. . . . .	31
3.19	Definition for function <b>Context()</b> part 2, typing. . . . .	32
3.20	Correctly identified faces (left) and skin regions that were sent to KNN (right). . . . .	33
3.21	Process of selecting a region for KNN and the output. . . . .	34
3.22	A flowchart representing the steps to identifying a face. . . . .	35
3.23	Pseudocode for function <b>Context()</b> part 3, talking. . . . .	35
4.1	Subfigures and display correct classification results for the color and context model of finding a pencil. . . . .	38

*List of Figures*

4.2	A correctly classified writing activity. . . . .	39
4.3	Results in an outside unseen video from the AOLME database. . . .	39
4.4	Results from model in an outside unseen video from the AOLME dataset. . . . .	40
4.5	Correct keyboard identification. . . . .	40
4.6	Correct classification in outside unseen video from AOLME dataset.	41
4.7	Correct classification of typing activity. . . . .	41
4.8	Correct classification of not typing in difficult case. . . . .	42
4.9	Correct classification of typing activity in difficult case. . . . .	42
4.10	Subfigures and display correct classification results for the color and correct KNN results for the face. . . . .	43
4.11	Correct classification of face KNN during validation. . . . .	43
4.12	Subfigures and display correct classification results for no talking and talking. . . . .	44
4.13	Correct classification of talking activity in difficult case in unseen video. . . . .	45
4.14	Incorrect classification of table and legs as keyboard. . . . .	46
4.15	Subfigures and display incorrect functionality of the face KNN from skin and blank space. . . . .	46
4.16	Classification of a mouse as a keyboard. . . . .	47
4.17	Table masking issue. . . . .	47

*List of Figures*

4.18	Incorrect masking due to whiteboard. . . . .	48
4.19	Keyboard classified as mouse. . . . .	48
4.20	Face missed due to hand on face. . . . .	49
4.21	Subfigures displaying incorrect classifications from all models. . . . .	50
4.22	Confusion matrices for object KNN. Mon stands for Monitor. . . . .	51
4.23	Confusion matrices for KNN Results. . . . .	52

# List of Tables

2.1	RGB+Optical Flow Methods . . . . .	8
2.2	Temporal RNN Methods . . . . .	10
2.3	Commonly Used Datasets . . . . .	11
3.1	Context Conditions for Object Recognition . . . . .	36
4.1	Dataset Video Types . . . . .	37
4.2	Dataset information for the final training dataset. . . . .	38
4.3	KNN Models for Typing and Talking . . . . .	51
4.4	KNN Results . . . . .	51
4.5	Tensor-flow KNN Results . . . . .	51
4.6	DNN Results . . . . .	52

# Glossary

AOLME	The Advancing Out of School Learning in Mathematics and Engineering research study.
SotA	State of the Art.
Binary Image	An image consisting only of black and white values.
Centroid	The center point in a contour.
Contour	The collection of points representing the outline of a 'blob' in a binary image.
CNN	Convolutional Neural Network.
DNN	Deep Neural Network.
Feature Space	The n-dimensional space in which the variables to be classified by a machine learning algorithm to be classified are stored.
Ground Truth	A set of labelled data that serves as a point of comparison for classification accuracy.
HSV	Hue, Saturation, Value color model.



## *Glossary*

KNN	K-nearest neighbors, a machine learning model that selects a classification for an object based on the distance between the object of interest and its k closest neighbors in a graph.
LSTM	Long Short-Term Memory, a block of a neural network that serves as a building component, i.e. short term memory which can extend over long periods of time when many are layered together.
LCRN	Long-term Recurrent Convolutional Networks.
ProcNet	A Procedural Neural Network.
RGB	Red, Green, Blue; referring to a color image frame in a video.
RNN	Recurrent Neural Network.
ROI	Region of Interest.
STEM	Science, Technology, Engineering and Mathematics.
Video in the Wild	Video recorded in an uncontrolled environment.
YCrCb	Y for Luminance, Cb for Chrominance-Blue, and Cr for Chrominance-Red. Another color value model.

# Chapter 1

## Introduction

### 1.1 Overview

Human activity recognition in raw video is a challenging problem that has drawn a lot of interest, especially around classification tasks involving huge video datasets. Most current methods have focused on speeding-up previous approaches, methods by which would otherwise require an impossible amount of human time to manually classify. However, these classification methods, even when assisted by humans, are not yet designed to accurately handle *video in the wild*.

An example of a large video database (several TB worth of raw video at the time of this writing) is the AOLME video database. The dataset collected for the AOLME project is an agglomeration of information in raw video that is currently being analyzed manually. The data were collected for the purposes of understanding how children learn in situations involving mathematical and programming challenges, such that teaching methods can be improved to broaden underrepresented student participation in STEM fields. The information necessary for identifying key features of the video include the discovery of interactions between the students and their

## *Chapter 1. Introduction*

facilitators for the purposes of identifying best teaching practices.

Because of the sheer size of the dataset, much of the video will never be fully analyzed. This being, in part, due to the many hours necessary to manually classify each video for each individual activity. In these cases, clearly an efficient, accurate classification tool to aid in the discovery for those researchers such that they could quickly identify videos bearing certain activities is warranted and could greatly improve the quality of their results.

## **1.2 Motivation**

The majority of current human activity detection methods require a significant amount of human time to clip out regions of interest and train a system to correctly identify actions within each region. Referred to as the ground truth, this manual labeling of cropped video segments has been used for training for the majority of machine learning classifiers to date. The current thesis focuses on identifying multiple video activities in raw videos without the requirement for manual cropping individual video segments prior to automatic classification.

Chapter 1. Introduction



(a)



(b)



(c)



(d)



(e)



(f)

Figure 1.1: Writing (a) v.s Not Writing (b), Typing (c) v.s. Not Typing (d), and Talking (e) v.s. Not Talking (f)

The thesis focuses on the detection of human video activities that are shown in Figure 1.1. More specifically, the thesis will consider: writing vs not writing as seen in Fig. 1.1a and 1.1b, typing versus not typing as seen in 1.1c and 1.1d, and finally talking versus not talking in 1.1e and 1.1f within a full, uncropped video. In comparison, more conventional methods would consider the same detection problems where the video activity has been isolated as demonstrated in Fig. 1.2. On the other hand, for training purposes, the thesis will consider training on the cropped video activities. To differentiate between the two database types, we refer to the original

videos as *raw videos* as opposed to the *cropped videos*.

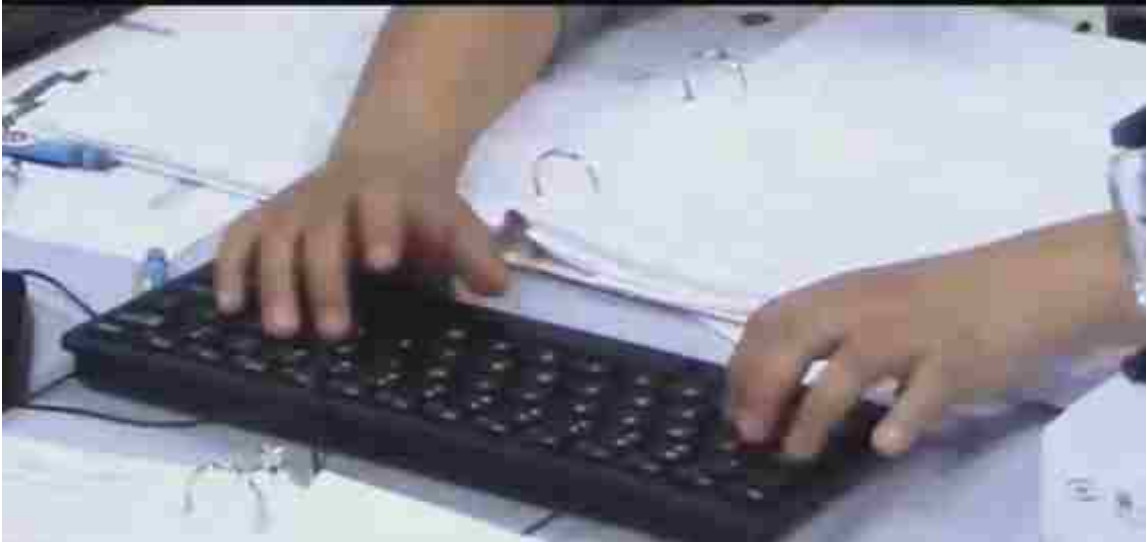


Figure 1.2: Cropped video sample.

The proposed research presents a novel approach to classification of video activity which can classify human activity with regards to the contextualized interaction of a collection of objects which have been identified using color. The method attempts to better classify human activity by considering the statistics of motion vectors extracted from regions of interest associated with adjacent objects. The approach is different from the standard methods that classify human activity through the use of full-frame methods that cannot be associated with specific objects or motions.

### 1.3 Thesis Statement

The thesis of this research states that a specific human activity of interest can be classified in raw video given a combined set of objects identified by means of color, motion vectors about the primary object, and the use of contextualization algorithms for each sought activity. The research of this thesis assigns a predicted classifica-

tion based on the interactions between the objects, referred to as giving context to an interaction. A final classification may then be assigned using machine learning techniques to examine motion vectors within the region of interest which has been identified. This thesis is focused on the determination that the proposed method which contextualizes object recognition prior to analyzing motion vectors is a viable solution for detecting specific individual activities in raw video, and additionally within other videos unseen by training. Furthermore, this thesis states the provided method is easily adjustable for identification of activity given various other human-object interactions by their associated context, and that activity recognition can still be ascertained given alternate methods of object detection.

## **1.4 Contributions**

The primary contribution of this thesis is to provide a new method for determining human activity based on the interactions of detected objects in raw videos. Mainly, the contribution is that the thesis provides a method for a dramatic reduction in the number of features in activity recognition. The thesis adds locality to the many RGB+flow methods that are currently considered among the most successful classification efforts, allowing context to be applied to the exact activity being sought. Human activities restricted to small regions of interest that were not previously classifiable by other methods can be analyzed with the proposed method (e.g., differentiating between a student playing with a pencil and writing).

## **1.5 Summary**

The remainder of this thesis is divided into five chapters. In Chapter 2 we explore the current state of the art methods for activity recognition in video, along with

## *Chapter 1. Introduction*

their respective datasets, and supply ample comparison to our algorithm. Chapter 3 explores the proposed algorithm in detail, beginning with an overview of the full model and breaking down the method to individual algorithms. A presentation of the findings are displayed visually, summarized into tables, and briefly discussed in Chapter 4. Finally, Chapter 5 continues with a summary discussion of the results and closes with conclusions and suggested future work. An appendix of code is included in Chapter 6 for reference.

# Chapter 2

## Background

Video activity classification is an expansive area of study which yearly generates an enormous amount of interest. Currently there is no generalizable solution, thus there have been multiple attempts to solve the problem in a way that can be applied outside of training datasets. Of particular interest are effective algorithms to segment video that has not been preprocessed in any way for classification of activity. This is made difficult by the numerous, transient human activities that may take place in any single video because they cannot be readily separated locally, and sometimes may be recognized as a motion within a sea of noise. Confounding issues further, with classified models trained on small clips of single activities of interest, an extremely computationally expensive comparison of the entire feature space of a set of vectors to match the trained template is necessary. Again, this fails in the case that an activity is associated with motions over small regions of interest, though the very existence of the sought activity could be ascertained with the human eye.

The state of the art explored in the literature review for this thesis focused heavily on methods of temporal frame differencing, bag of features, RNNs to recursively learn annotated video, optical flow methods for motion estimation, CNNs for feature



Study	Dataset	Comments
Two-Stream Convolutional Networks for Action Recognition in Videos, <b>2014</b> [274]	UCF-101 and HMDB-51	Combines RGB frame and flow inputs for frame by frame recognition on a new network combination. 87.0% on UFC-101, 59.4% HMDB-51
Beyond Short Snippets: Deep Networks for Video Classification, <b>2015</b> [298]	Sports 1 million dataset, and the UCF-101 datasets with and without additional optical flow information	Improved dense trajectory motion estimation by removing camera movement by (73.1% vs. 60.9%) ,(88.6% vs. 88.0%,And (82.6% vs. 73.0%), respectively
Long-term Recurrent Convolutional Networks for Visual Recognition and Description, <b>2016</b> [56]	UCF101 dataset, AlexNet, CaffeNet	LCRN builds on [274], applies models to conventional activity challenges. Averages all frames to classify, RGB: 68.20%, Flow: 77.28%
Real-time Action Recognition with Enhanced Motion Vector CNNs, <b>2016</b> [299]	UFC101, ImageNet and ILSVRC-2012 datasets	Aim to speed up the RGB-Optical flow method [274] for use with realtime video classification 61.5% 403.2fps

Table 2.1: RGB+Optical Flow Methods

extraction from a single frame along with frame differencing for motion (flow), or some combination of the above. A summary of current research is provided in Table 2.1. All methods pertain to frame feature classification and optical flow.

Many of the attempts to classify video actions are now using two or more trained neural networks to make their predictions. The methods pass each frame for feature extraction through a CNN then form a list of objects, combine with motion estimated

## Chapter 2. Background

through an RNN or other type of recursive learning network, typically an LSTM method, then combine the two outputs to make a prediction for the activity based on learning models or probabilities. One such attempt passed a list of all the activities the video contained along with the video to train a weakly supervised RNN, making use of youtube annotations [261].

A method similar to the proposed method feeds RGB data to one trained network and optical flow data to another finally combining them with an LSTM achieving 87% accuracy on the UFC101 dataset [274]. This research is used as the basis for many of the following bodies of research. They allow a CNN to classify all objects in training frames prior to sending the motion vectors for classification, then after having a classification for every frame give the action a classification.

An example study that used different combinations of neural networks for this type of classification, such as LSTM instead of RNN, was done by [298] and barely improved upon the previous method [274]. Their model attempted to classify a video with only raw frames as well. Their results are interesting to us as when they removed optical flow they saw only a minor decrease in accuracy which could indicate either overfitting or the probability models being independent from one another.

More recently a study that builds from [274] simply attempted to speed up the framerate so that it could be applied in realtime as opposed to waiting for all frames to be estimated [299]. Some of these models such as in [298] sought ways to improve the motion vector classification, also seen in [291]. We begin to see a pattern in the state of the art methods where followup methods do not attempt to compete but to improve upon the base RGB+flow method.

Methods developed within the last year are looking more toward the use of LSTM to identify sequences of frames over short time-spans. These are referred to mostly as temporal frame extraction techniques with segment procedural learning and require

Study	Dataset	Comments
Recurrent Assistance: Cross-Dataset Training of LSTMs on Kitchen Tasks, <b>2017</b> [249]	Breakfast (18.8), 50 Salads (46.86%), and MPII Cooking 2 (41.1%)	68.5% and 94.0% on HMDB51 and UCF101 respectively, 18.59% average obtained by the (best performing) RGB stream in their model
Towards Automatic Learning of Procedures from Web Instructional Videos, <b>2017</b> [300]	A series of instructional cooking videos	Segment/label a video and train a ProcNet to recognize each frame based on previous labels; 30.4% precision, 37.1% recall
Weakly Supervised Action Learning with RNN based Fine-to-coarse Modeling, <b>2017</b> [261]	Breakfast dataset and Hollywood extended dataset	Provide a list of activities and allow RNN to classify each activity using feature extraction Accuracy 33.3%, SotA: 27.7%

Table 2.2: Temporal RNN Methods

an incredible amount of time training machines or large annotated youtube datasets which can be referenced quickly from Table 2.3. We summarize these works in Table 2.2.

One such study provided an already labeled activity and used an RNN to learn the activity using objects that were found in the frame, somewhat of a reverse of the previous case [261]. The paper gave an example of a hand, teabag, and teacup indicated somebody was making tea. Another method used what they referred to as ProcNets, or procedural neural networks, as a weakly supervised approach to learning based on temporal alignment [300]. Here, they use a similar method which used previous information to determine if the estimate "made sense". The example given was that it isn't likely that someone is mixing dough if they cracked an egg

Chapter 2. Background

and then two frames later are cracking an egg [249]. Several more studies of LSTM only approaches both supervised and unsupervised have been reported in [279] and [137].

<b>Study</b>	<b>Dataset</b>	<b>Comments</b>
Hollywood Extended Dataset	937 videos with 16 different actions from 69 movies.	<a href="#">Hollywood Extended Dataset</a>
MPII Cooking 2	87 unique cooking activities with annotations .	<a href="#">MPII Cooking 2</a>
Breakfast Dataset	A video dataset containing video of 10 common breakfast chores along with annotation.	<a href="#">Breakfast Dataset</a>
UCF-101	101 human action classes with over 13,000 video clips.	<a href="#">UCF101 Dataset</a>
HMDB-51	A human motion dataset consisting of 51 human actions.	<a href="#">HMDB-51 Dataset</a>
ImageNet	A large image database specifically for image software research projects.	<a href="#">ImageNet</a>
YouTube-8M	YouTube video IDs and labels from 4700+ entities.	<a href="#">YouTube-8M</a>
YouCook	88 cooking videos from youtube.	<a href="#">YouCook II</a>

Table 2.3: Commonly Used Datasets

The classic approach to training consists of segmenting videos into short clips which are pre-processed, labeled, and then used as input to a machine learning model. The fit data is then tested using other clips showing the same activity. These video clips contain typically only one person performing one action, or are sent after a

## *Chapter 2. Background*

machine has been trained with a series of labeled images; there is a massive amount of object training that must be done in either situation. Training a cascade to recognize a new object is cumbersome and can take quite a bit of time, and sometimes will not be enough to recognize the object of interest. Imagenet and other repositories have readily available feature cascades model zoos which are quite limited in terms of content, or objects, they are trained to recognize; though, they are rapidly becoming more diverse. Currently trained object recognition cascades or machines are few and far between, sometimes not publicly available, and do not necessarily detect objects related to education, such as a pencil, pen, eraser, or other such writing implements. These objects are of particular interest in the AOLME video library used in the testing of this model, so having a way to detect writing implements accurately via object detection would have been of great value to this research.

Further, review discovered that most all studies used the datasets presented in Table 2.3 for benchmarking the performance of each presented algorithm. These datasets are used as points of comparison for classification typically of objects of different human activities. Though these are the common datasets used, they do not have models of all objects under review, nor the particular activities being explored, and they do not provide the scale of data necessary to perform a robust review of the algorithm presented. The AOLME dataset contains several hours of human activity in the wild, rather than zoomed on a single action, and have many examples of the activities we wish to identify. This dataset has an inconsistent number of actions happening per frame, all of which need to be assigned context in a manner that the above datasets do not lend well to. Finally, the AOLME dataset has a consistent color scheme to the data provided (the objects used are all the same make/model), and as discussed, this algorithm relies on color to find objects, so for this thesis, a subset of this video database will be used to test the performance of the algorithm.

The AOLME project is an interdisciplinary collaboration between the UNM Col-

lege of Education and the School of Engineering (see [41]). Some of the earlier lessons learned in the AOLME project have been summarized in [247]. The AOLME dataset generated a large video dataset that presents several computing challenges (e.g., see similar problems in [197]). In research related to the current thesis, Cody Eilar looked at writing and typing classification using Amazon Web Services [57]. The recent emergence of convolutional neural networks is also of great interest to the ivPCL lab at UNM because of a long tradition in convolution-based methods (e.g., see [40], [37]) and image analysis methods (e.g., see [225], [237], [164], [156]).

## 2.1 Classification of Human Activity in Video

In review, many current ideas for activity recognition rely heavily on optical flow motion vectors to determine what is happening in the video clip. In a large video, where many activities could simultaneously be taking place, this creates an obvious problem of structural noise in the motion vectors, the obscuration of which causes extreme difficulty when attempting to hone in on a specific motion. Additionally, calculating the feature vectors for an entire video is highly computationally expensive, and in attaining a good set of ground truth data often many different videos are required.

These issues lend to complicated neural networks being used to quickly facilitate the learning process. Neural Networks can have significant fallbacks as they are very prone to overfitting and in many cases perform as "block boxes", we simply cannot tell how they are coming to their conclusion or how they fail. This phenomenon is referred to as "unexplainable AI". In such systems, the problem is that the systems come to conclusions without being able to explain how the results were achieved. This, of course, is a significant problem.

Though some of the studies we reviewed inspected various approaches to overcome

## Chapter 2. Background

these limitations (Temporal learning and speeding up of flow methods for realtime classification), they still suffer serious fallbacks. When applied to outside datasets, these methods tend to fail and do not generalize outside of their training dataset. Though our method also makes use of a neural network, it was not necessary to use a CNN or RNN to identify objects then learn features. Additionally, the features in this thesis are made very easy to visualize such that it is possible to inspect exactly where and how the algorithm fails.

There are many obstacles preventing generalization of activity recognition methods. High, structural noise motions, as previously mentioned, such as walking will occlude lower magnitude motions such as writing during background subtraction. This creates a problem of bleeding when creating blobs about which we are attempting to classify; the feature space then is nearly impossible to segment in a way such that each individual motion can be found, much less classified. This is further complicated by illumination variations, which can cause motion to be found in a video where there is no motion. This also has a great effect on motion vector analysis as high noise motions can drown out smaller motions. Geometric distortion and locality of an action in a full frame can also cause significant difficulty with classification. There are image techniques which classify a image's scenet based on the relation of various objects found in its feature space [110]. We draw from this idea and attempt to generalize it for use in video by coming up with a geometrically invariant object interaction algorithm based on color. The proposed method removes the difficulty in determining an action's existence in different datasets simply by first determining if an object of interest is present within a certain context. This may be a combination of objects, such as a plate, a fork, and an up and down motion, or simply an object itself, e.g. a pencil, being moved around by a human (which is made as an assumption since pencils cannot move on their own!)

More recently, researchers have been combining depth cameras to help identify

## *Chapter 2. Background*

primary motion and colored object recognition. Since the AOLME dataset was recorded on a normal camera, we could not employ these methods and thus they were not pursued further. With regards to object recognition, this is currently done by using pre-trained CNNs or model zoos, such as AlexNet, to identify and classify features within a single frame. This is an exhaustively explored method of study with excellent reported accuracy on common datasets. Here, we note that we choose to implement the least computationally expensive method, recognition of objects using color. For the purposes of this research, the objects we wish to classify are primarily of uniform color. This thesis then follows a similar pattern to the more successful methods in that we will be combining object recognition with optical flow, though in a greatly reduced feature space.



# Chapter 3

## Methodology

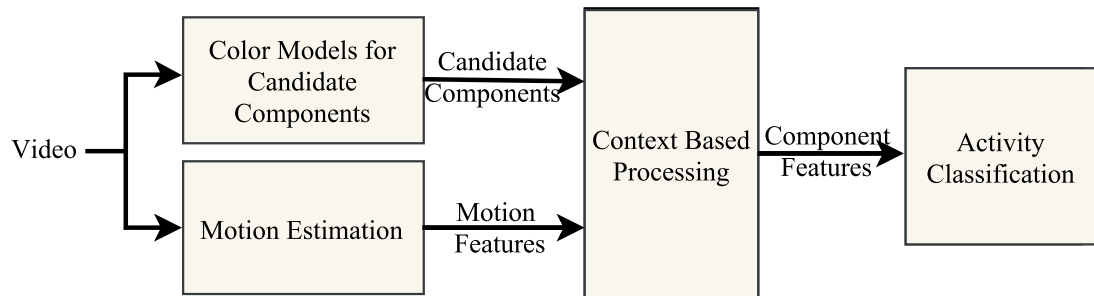


Figure 3.1: Overview flowchart.

### 3.1 Overview of the Approach

We present an overview of the method in 3.1. Initially, we apply color-based segmentation to extract the candidate objects of interest. For each video frame, we extract motion vectors that are specific to the objects of interest. We then apply context-based rules to filter and identify components that can be associated with specific activities. For the specific components of interest, we extract motion vector

features that are used for activity classification. In what follows, we describe the various methods that are involved.

## 3.2 Candidate Region Selection Based on Color Models

### 3.2.1 Color Models for Pencils, Table, Paper, Keyboard, and Faces

The goal of our use of the color models is to determine candidate components that are further processed based on their relative context. Thus, our use of color models produces an over-segmentation of the objects of interest. We will later correct the over-segmentation result by eliminating invalid components based on context.

We develop a color-based segmentation system based on the HSV color space. Since we are working with single-color objects, the goal was to determine bounds for each component. Bound selection was done visually using a simple database of 20 examples of pencils, paper, tables, skin regions and keyboards. Here, we found thresholds that worked on all images at the same time, as verified in the visual display of all examples. An example is shown in Fig. 3.2. The code for finding pencils can be found in Appendix A, specifically under A.1, and A.2.

Caution was taken to first visualize the selected colors in the video to confirm accurate detection. A ROI box was drawn on the frame around each centroid of each contour that remained after masking. Satisfied with a visual inspection over various videos from different datasets, we did not further optimize the values for colorspace. An example of the visual inspection is seen in Fig. 3.4. A flowchart summary of code is provided in Fig. 3.5.



Figure 3.2: Masked pencil color (right) v.s. original frame (left).



Collection of Pencils

Binary Version

Figure 3.3: Subfigure a shows the collection of pencil variations we see, and b shows what we wish to achieve after masking and thresholding.



Figure 3.4: A cropped writing video frame with potential pencil highlighted.

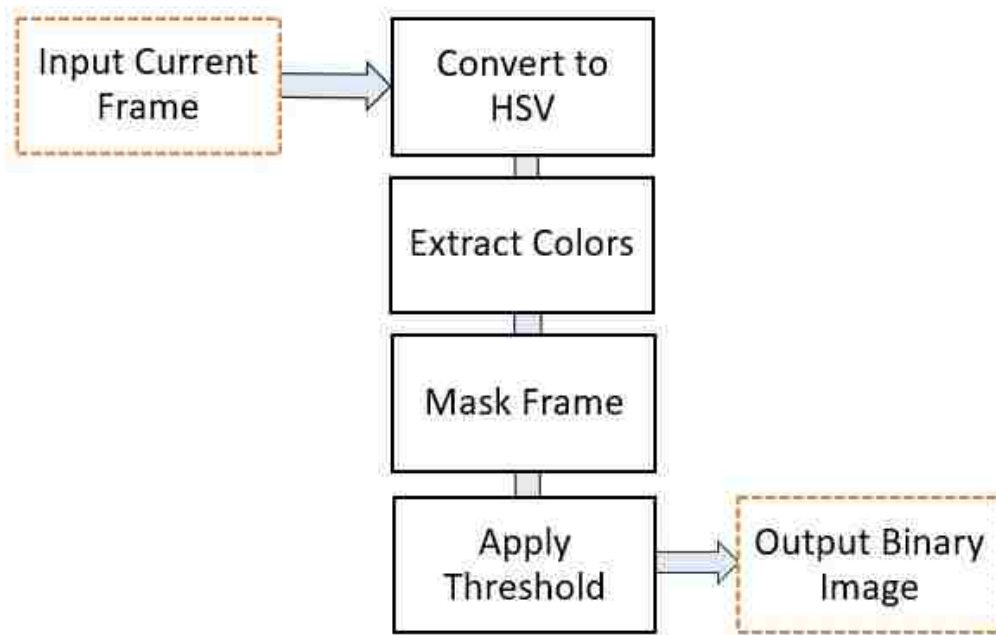


Figure 3.5: Flowchart displaying the process of extracting colors from a frame.

## 3.3 Context-based Processing and Feature Extraction

The color models provide candidate regions for further processing. The candidate regions need to be carefully selected and then processed for context by processing relations between them. Then, a combination of checks is applied to check for interactions between them and motion content. The histograms of motion magnitudes and orientations are then used as features for further classification. In what follows, we provide details for each step.

### 3.3.1 KNN Classifier for Selecting Keyboard and Face Components

For keyboard and face detection, our use of the color models provides for initial candidate regions. Here, we found that these initial candidate regions needed to be further processed prior to use. We apply morphological filtering to remove minor regions. We then compute the bounding box for each region, zero-pad, convert to grayscale (Y component in Y-Cr-Cb), and use bilinear interpolation to resize each one of them to 128x128. We then use K-nearest neighbor to classify each component (e.g., keyboard present or not). For the pre-KNN code, refer to [A.3](#) for the keyboard, [A.4](#) for the table, and [A.5](#) for the skin detection. Figures [3.6](#) and [3.7](#) contain several examples of data used in the KNN training models.

The faces KNN classifier is trained using a collection of 1,700 images from various videos within our datasets, and labeled 0 or 1 to identify faces versus not faces respectively. For the keyboard, mouse, and monitor, we used 170 images. Refer to section [3.4](#).

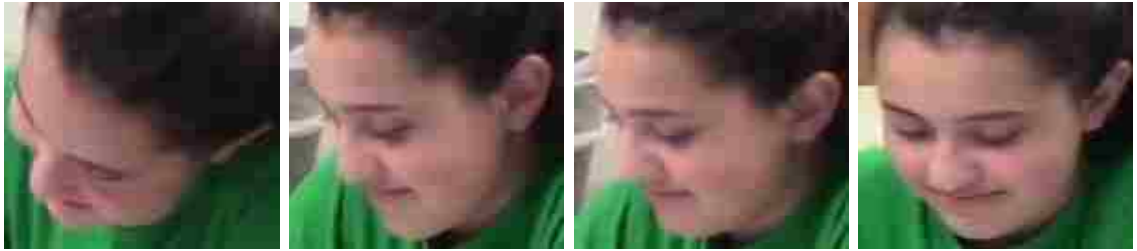


Figure 3.6: A sample of faces that are used to train the face KNN model.

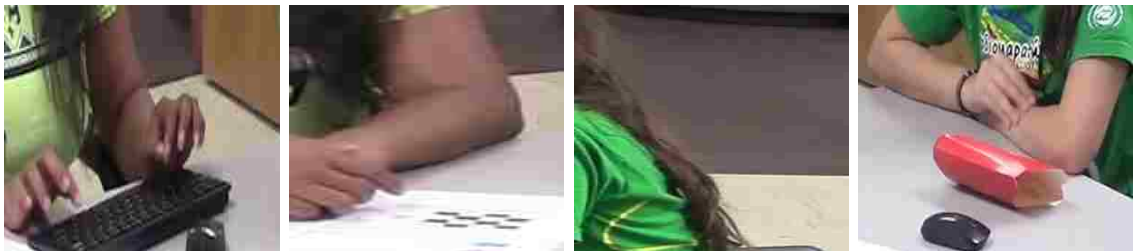


Figure 3.7: A collection of "Non-Faces" used in training the face KNN model.

### 3.3.2 Writing

We summarize context-based feature extraction for the writing activity in flowchart format in Fig. 3.13 and pseudocode format in Fig. 3.14. In what follows, we describe each step.

For writing, we first use the binary image created by masking the table and take the bitwise and between the images to preserve only the objects on the table. This mask is applied to both the paper and the pencil binary images.

The following step uses the binary image of the pencil as input to a find contours function in opencv, after which we use the moment of each contour and extract each centroid, and after obtaining the best fitting rectangle, calculate the aspect ratio of the contour. We check to make sure the aspect ratio and area are appropriate for a pencil, and then finally, using the centroid, we create a region of interest box around the pencil.

### *Chapter 3. Methodology*

We use the dimensions of the region of interest box plus some padding to slice the paper binary image, effectively cutting out only areas near the pencil, and the sum of all the pixels remaining are taken. If the sum is greater than zero in this area, we determine that the pencil is on top of a table and over top of a paper, thus writing may be present.





Figure 3.8: A collection of figures used in training the keyboard KNN, with keyboard samples in (3.8a), (3.8b), and (3.8c), and non-keyboards (monitors in (3.8d), (3.8e) and mice in (3.8f), (3.8g), (3.8h))

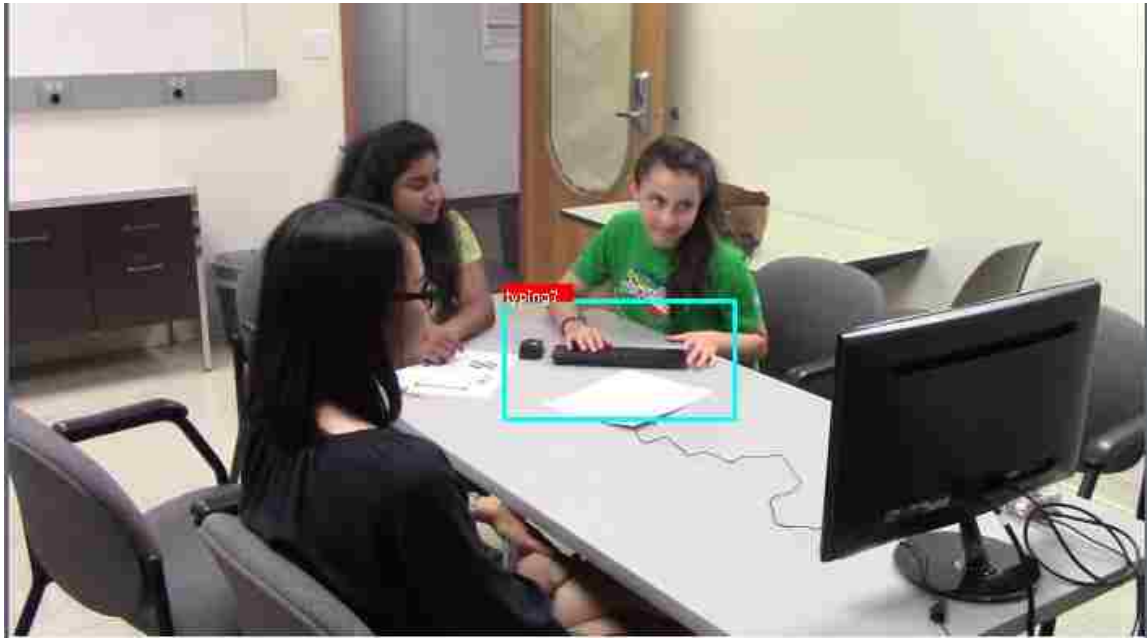


Figure 3.9: A correctly identified keyboard using KNN

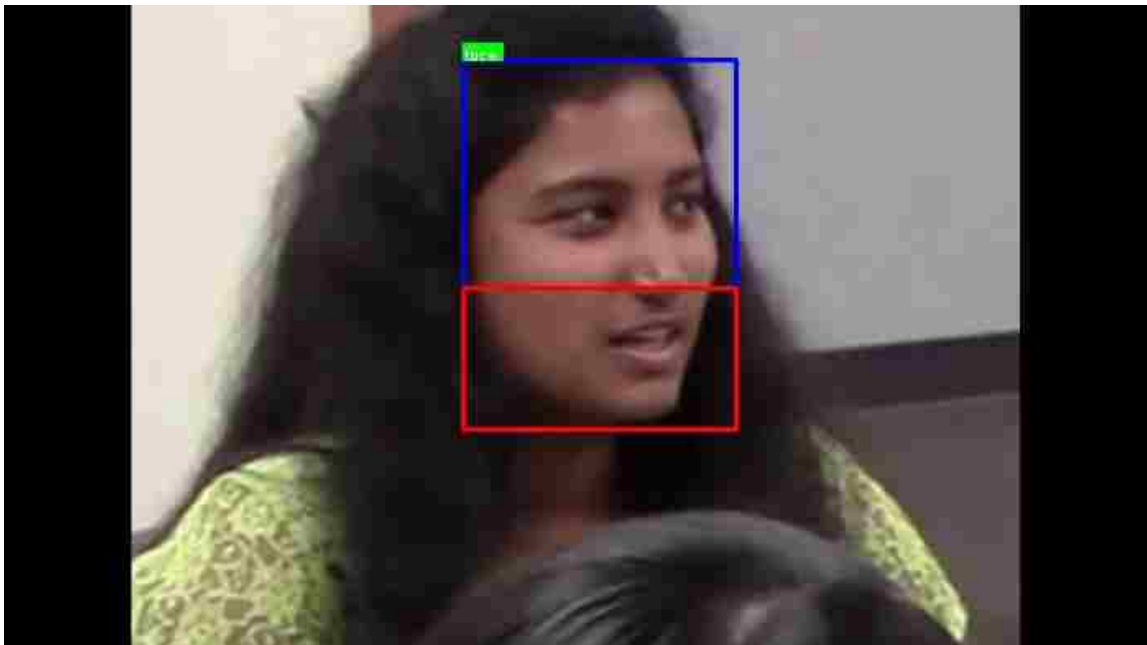
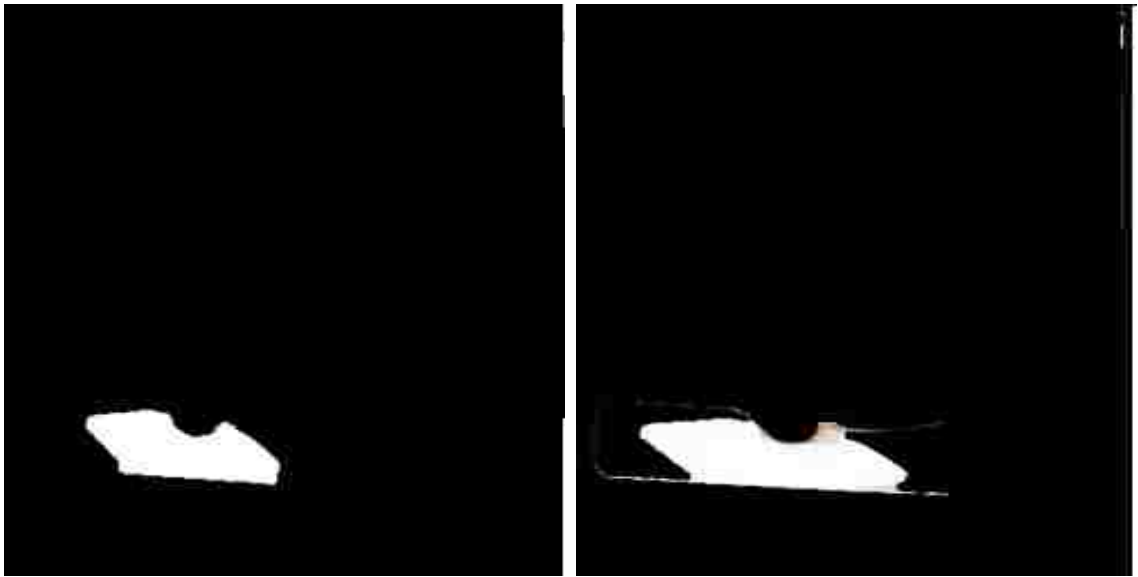


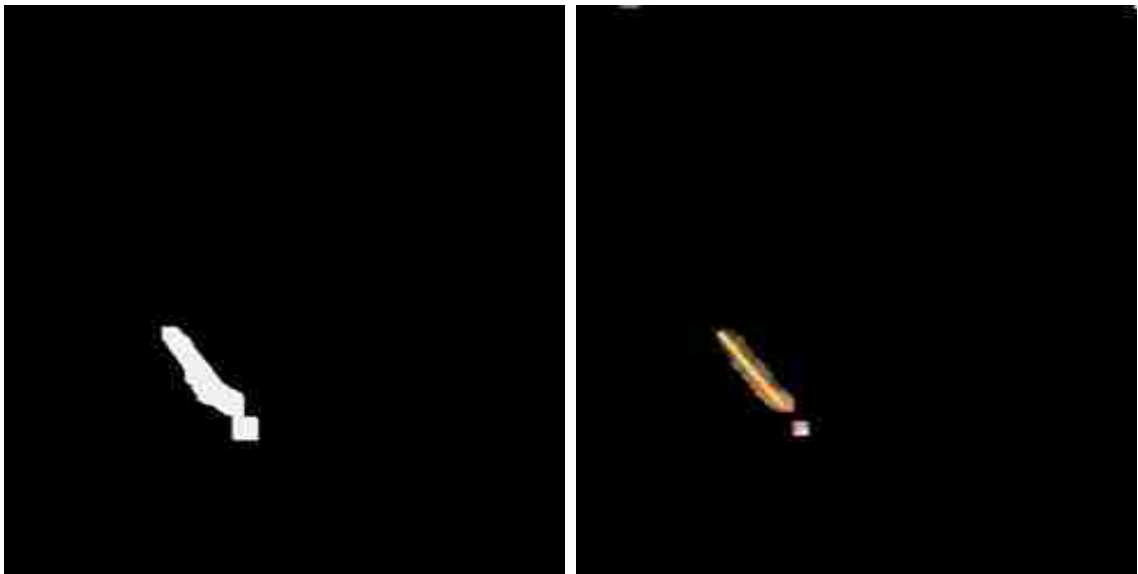
Figure 3.10: A correctly identified face in KNN.



(a) Binary Paper Image

(b) Masked Paper Image

Figure 3.11: Paper detection by color, intermediate steps.



(a) Pencil Binary Image

(b) Masked Pencil Image

Figure 3.12: Pencil detection by color, intermediate steps.

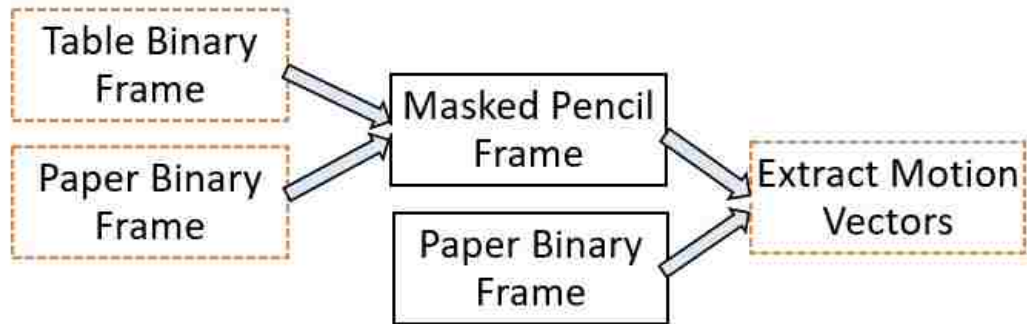


Figure 3.13: Flowchart for confirmation of a pencil's presence prior to motion vector extraction.

The final step is to check for motion within the region of interest using the flow passed to the contextualization function. If a certain motion magnitude threshold is exceeded, we then extract the motion vectors from this region of interest and calculate a histogram from them, which is returned by the function and later used to determine whether the motion is due to writing or not.

Parameter optimization was performed through a visual interface. Here, the effect of the use of different parameter values is assessed through their impact on the ROI bounding boxes on a small set of training videos with different geometric angles, lighting conditions, and settings.

---

```

1 function Context (frame, obj1, obj2, flow, obj_names, centers, hist_list);
2 if Looking for Pencil, Paper then
3   pencil_image ← Bitwise And Binary Pencil and Binary Table
4   for All Contours do
5     Check Requirements
6     if Requirements are Met then
7       centroids ← Search for nearby Centroid
8       if No Centroid Nearby then
9         centroids ← Append Centroid
10      end
11      if Paper is Below Pencil then
12        if Motion is Found then
13          hist_list ← Calculate and Append Histograms
14          frame ← Draw Rectangles
15        end
16      end
17    end
18  end
19  return frame, hist_list, centroids
20 end
21

```

▷ Function continues...

---

Figure 3.14: Pseudocode for function **Context()** part 1, writing.

### 3.3.3 Typing

We provide a flowchart summary of the typing activity algorithm in Fig. 3.18 and provide pseudocode for the typing portion of function **Context()** in Fig. 3.19. In what follows, we explain each step.

Typing follows a similar identification method as for pencil detection, except for the need for an extra step needed to eliminate large gaps between the keyboard and the monitor. To overcome this, we calculate the convex hull from the contour of the table, and apply a bitwise and with the threshold image the hull is applied to in

order to close the holes within the table.

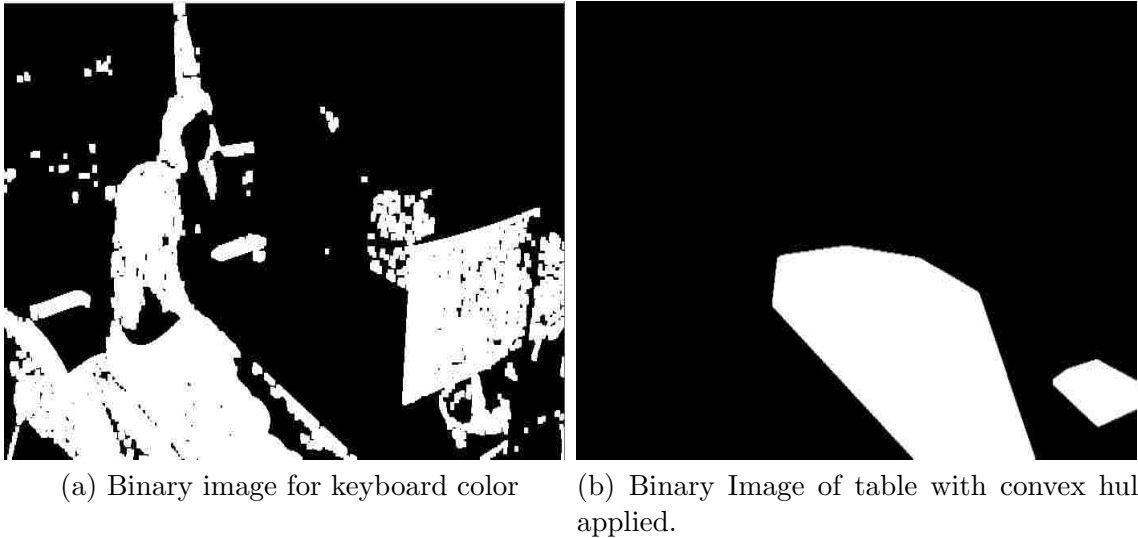


Figure 3.15: Images used as input for masking the background of the keyboard image

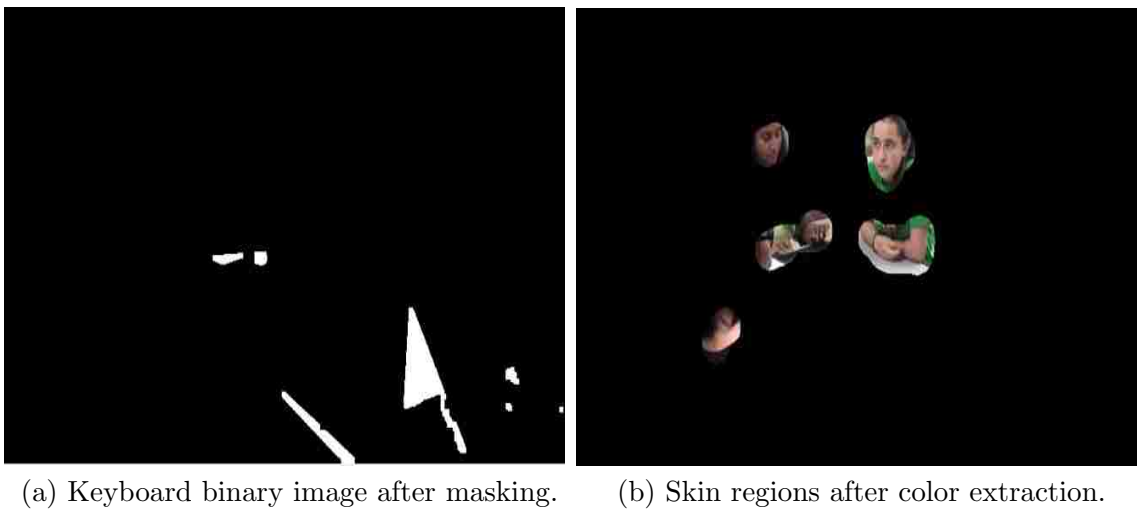


Figure 3.16: Figures (a) and (b) display the two components required to determine if typing may be present.

The contours are found for each of the objects in the keyboard threshold, and after the centroids are located, we use the ROI to first resize the slice to 128x128

### Chapter 3. Methodology

and fit the cropped grayscale image to the KNN model trained to classify keyboards, monitors and mice.

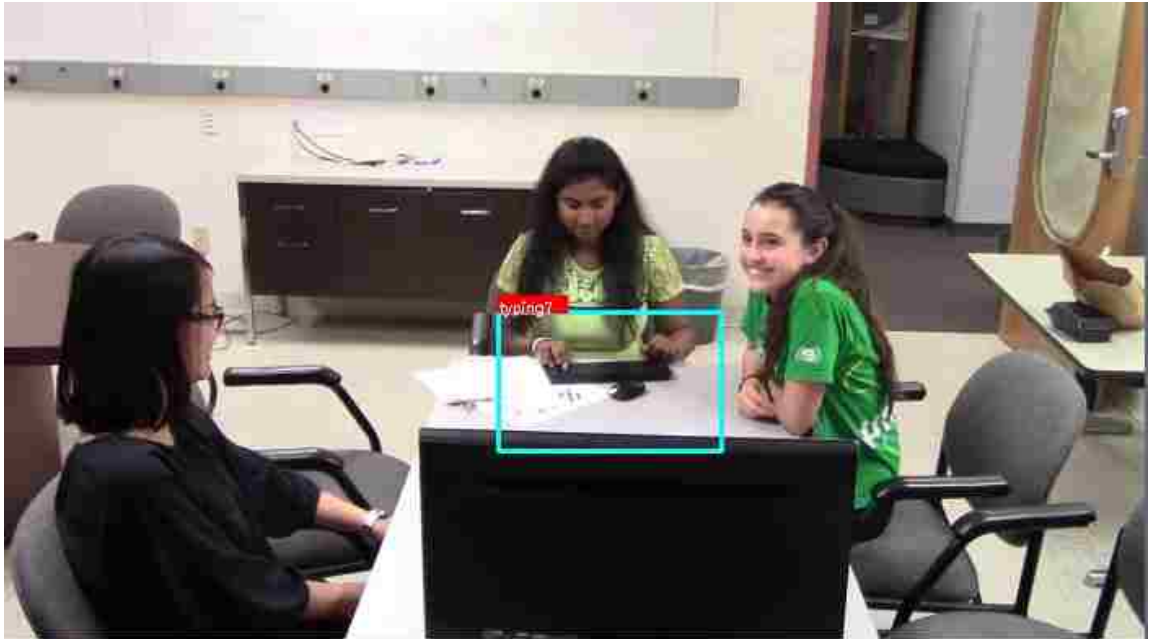


Figure 3.17: A keyboard which has been properly found and classified for motion vector extraction.

If the object is determined to be a keyboard by the model, we then do a check for hands by slicing the skin image within the ROI box and checking to see if the remaining pixel value sum is greater than zero. If it is, we then calculate the appended histogram of the magnitude and phase of the optical flow as done previously. Again we visually review the resulting ROI boxes as an informal optimization of the parameters of each of our contexts.

Chapter 3. Methodology

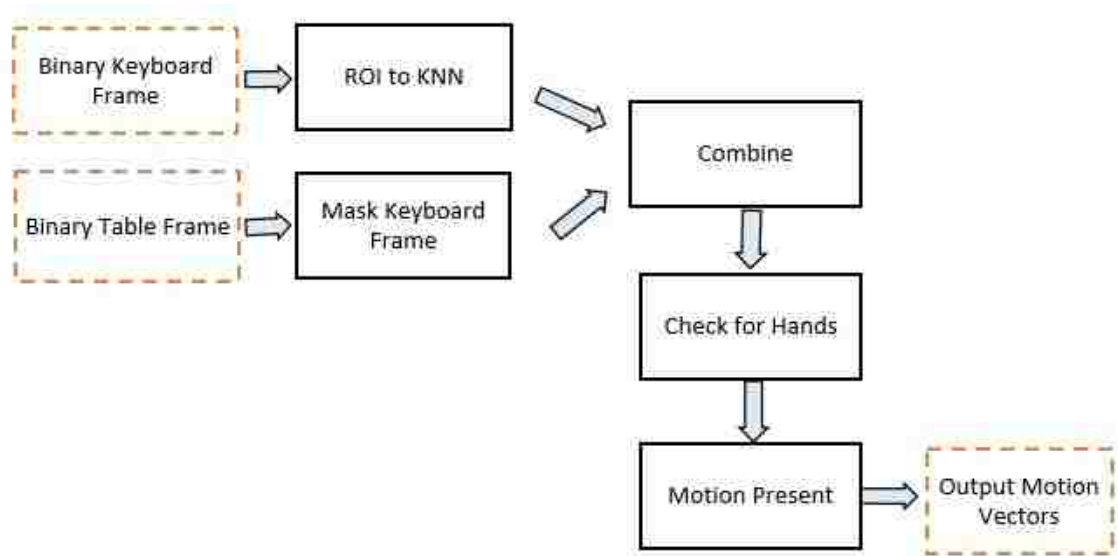


Figure 3.18: A flowchart representing the steps to extracting the keyboard and discovery of potential typing.



---

```

1 if Looking for "Keyboard, Table" then
2   table ← Calculate Convex Hull of Table Binary Image
3   contours ← Bitwise And Table and Keyboard Binary Images
4   for All Contours do
5     Check Requirements
6     if Requirements are Met then
7       centroids ← Search for nearby Centroid
8       if No Centroid Nearby then
9         centroids ← Append Centroid
10      end
11      keyboard ← Fit ROI to KNN
12      if Keyboard is Found then
13        if Hands over Keyboard then
14          if Motion is Found then
15            hist_list ← Calculate and Append Histograms
16            frame ← Draw Rectangles
17          end
18        end
19      end
20    end
21  end
22  return frame, hist_list, centroids
23 end
24

```

▷ Function continues...

---

Figure 3.19: Definition for function **Context()** part 2, typing.

### 3.3.4 Talking

We provide a flowchart for detecting the talking activity in Fig. 3.22 and provide pseudocode in Fig. 3.23. In what follows, we summarize each step.

To identify talking, the steps are slightly different. The skin is found using the same technique as the pencil, first converting the frame to YCrCb space and extract the range which preserves only areas of color matching the defined minimum

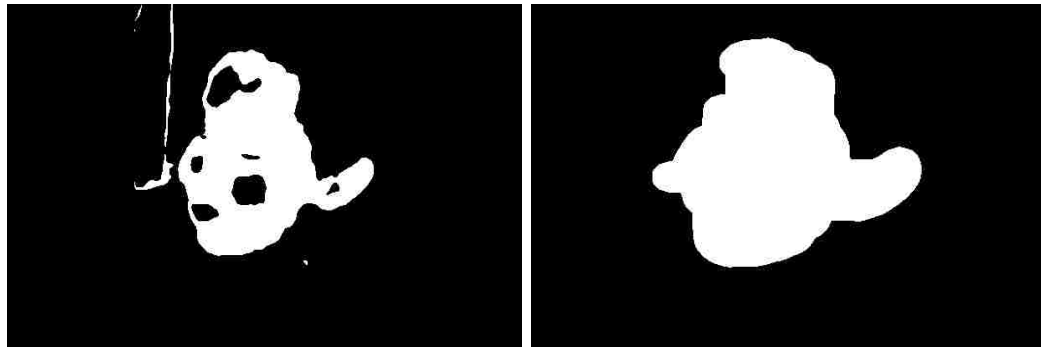
### Chapter 3. Methodology

and maximum. Some additional processing is done in this case to expand the areas around the extracted region using a gaussian blur before masking the frame, this includes also using a morphological open operation to remove noise from the similar colored bookshelves in the background. The blurring extrapolates a larger portion of the image around the skin area so the contour for evaluation in the KNN is more likely to contain the full face. Code is provided in appendix for the person object.



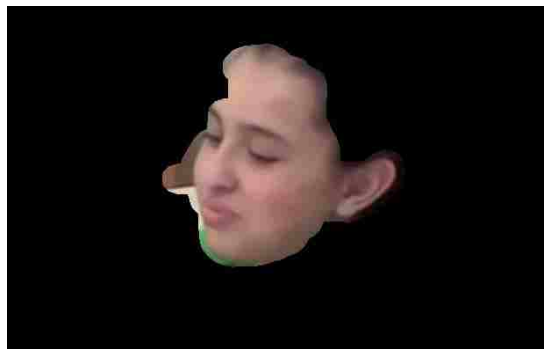
Figure 3.20: Correctly identified faces (left) and skin regions that were sent to KNN (right).

Using the threshold of this masked image, each contour is extracted and the boxed region about the centroid is resized to 128x128, as with the keyboard, and then passed to the KNN model. If the area is matched to being a face, we apply the golden ratio and check only the bottom portion of the face by dividing the height by 1.618 and adding to the y value to get the top y-location of the box. The golden ratio is commonly used in art and design applications, and is not often used in engineering practice. Application of this ratio adequately captures the mouth at any angle we are looking for. This area is checked for motion, and as the other methods, the motion vectors are calculated into a 1d histogram.



(a) Initial skin region mask.

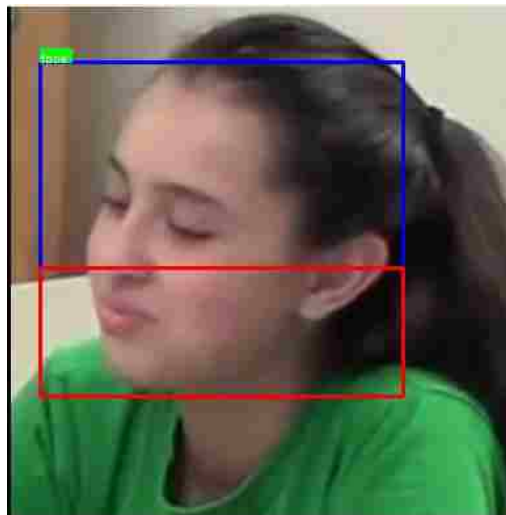
(b) Mask after filling holes.



(c) Mask applied to skin region.



(d) Area of image that is selected by the region.



(e) Resulting KNN output.

Figure 3.21: Process of selecting a region for KNN and the output.

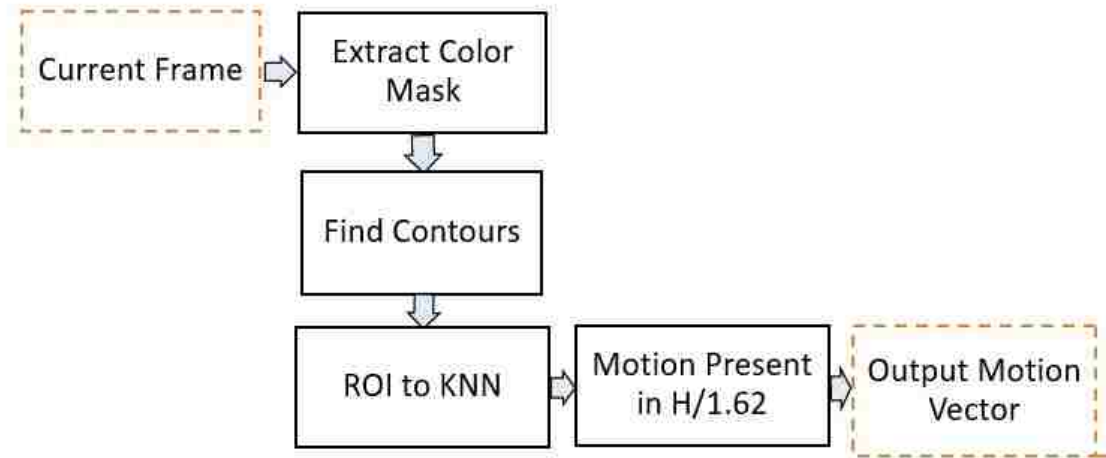


Figure 3.22: A flowchart representing the steps to identifying a face.

---

```

1 if Looking for Talking then
2   person_object ← PersonDetect()
3   for All Contours do
4     Check Requirements
5     if Requirements are Met then
6       centroids ← Search for Nearby Centroid
7       if No Centroid Nearby then
8         centroids ← Append Centroid
9       end
10      face ← Fit ROI with KNN
11      if face is found then
12        if motion found in  $\frac{face\_height}{1.618}$  then
13          hist_list ← Calculate and Append Histograms
14          frame ← Draw Rectangles
15        end
16      end
17    end
18  end
19  return frame, hist_list, centroids
20 end
  
```

---

Figure 3.23: Pseudocode for function **Context()** part 3, talking.

### 3.3.5 Feature vectors

The feature vectors are extracted every 3 frames as given in Table 3.1. Specifically, we calculate probability density functions (PDFs) of the magnitude and angle from the motion vectors and append them together. Each set of PDFs are separated by the centroid coordinates. Here, we note that we have a fixed size feature vector for each object. If a feature vector is missing, then the assumption is that the corresponding object is missing.

## 3.4 Classification

We classify feature vectors for object presence. We investigated the use of KNN and the use of fully-connected deep neural nets (DNN). Different classifiers were considered as described in the results.

	<b>Writing</b>	<b>Typing</b>	<b>Talking</b>
Inputs	Table A.4, Paper A.1, Pencil A.2	Table A.4, Skin A.5, Keyboard A.3	N/A
Object Check	$40 < \text{area} < 2000$ and aspect ratio $> 1.2$ or $< 0.5$	$200 < \text{area} < 8000$ and Has 3+ corners	$\text{area} > 1000$ and $0.5 \geq \text{aspect ratio} \geq 1.5$
Context Check 1	Pencil within table region	Keyboard detected on table	Face is detected
Context Check 2	Pencil near a piece of paper	Hands inside keyboard bounding box	Bottom part of face has motion
KNN Used	Yes	Yes	Color Only

Table 3.1: Context Conditions for Object Recognition

# Chapter 4

## Results

### 4.1 Dataset

We summarize the dataset types in Table 4.1 and the size of the dataset in Table 4.2. For model selection and proper reporting of the results, we used nested cross-validation. We report our final results using tenfold cross validation. Within the training set, we perform five-fold randomized cross-validation for model selection. For KNN, we investigated  $K = 3$  to  $K = 19$ , the use of both Euclidean and city-block distance for determining the nearest neighbor, and the use of feature scaling.

		<b>Writing</b>	<b>Typing</b>	<b>Talking</b>
Cropped	Train-	Positives Only	None	Positives and Neg-
ing	ing			atives
Video	Video			
Raw	Training	Negatives Only	Yes	Some Negatives
Video	Video			

Table 4.1: Dataset Video Types

	<b>Talking</b>	<b>Typing</b>	<b>Writing</b>
No. of Features	1755	1050	620
No. of Videos	14	14	15
FPS	60	26-60	24-60
Video Duration Range (in Seconds)	5-16	5-24	1-39

Table 4.2: Dataset information for the final training dataset.

## 4.2 Classification Model Results

### 4.2.1 Results for Writing

We present two accurate pencil identifications in Fig 4.1, showing that our color + context model is working correctly in these training cases.

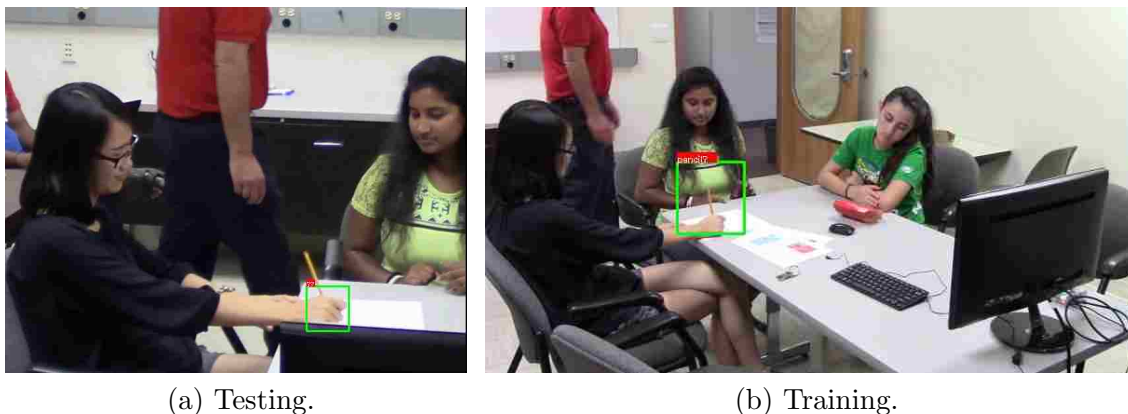


Figure 4.1: Subfigures and display correct classification results for the color and context model of finding a pencil.

We present an example of the KNN correctly classifying writing within a frame in Fig 4.2. Additionally we present an example of not- writing correctly identified in Fig 4.4 in unseen video from the AOLME dataset. These images are cropped or censored for privacy of the study participants.



Figure 4.2: A correctly classified writing activity.

For writing we achieved a highest accuracy using the DNN model with an approximate percentage of 92.5% accuracy. The dataset results are above in Table 4.2, and the remaining results for each classifier are found in Tables 4.3 thru 4.6.



Figure 4.3: Results in an outside unseen video from the AOLME database.





Figure 4.4: Results from model in an outside unseen video from the AOLME dataset.

Pencil not writing is correctly classified as not writing in Fig 4.4 with noise correctly classified as not writing; however, misclassified as being a pencil.

## 4.2.2 Results for Typing



Figure 4.5: Correct keyboard identification.

In Fig 4.5 we see correct identification of a keyboard using the color + KNN + context model.



Figure 4.6: Correct classification in outside unseen video from AOLME dataset.



Figure 4.7: Correct classification of typing activity.

Figs 4.6 and 4.7 show the correct classification of typing activity with our model, with Fig 4.6 being taken from unseen video from the AOLME dataset, and Fig 4.7 coming from our validation dataset.



Figure 4.8: Correct classification of not typing in difficult case.

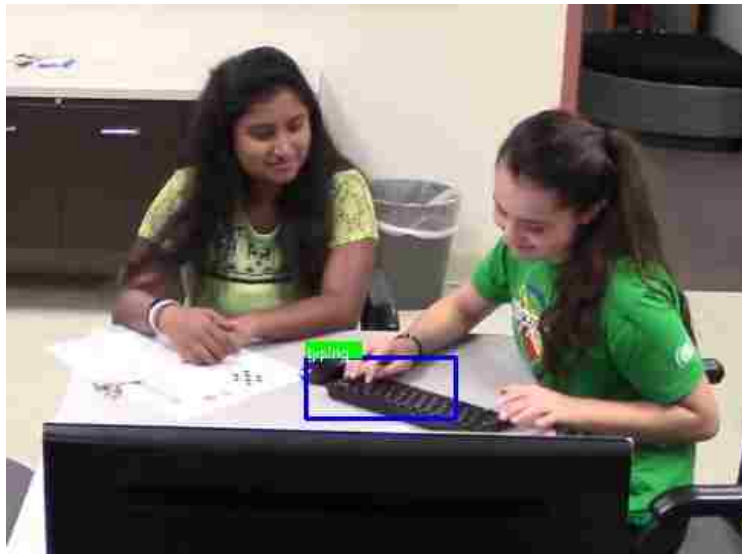


Figure 4.9: Correct classification of typing activity in difficult case.

The figures above display correct classification of no typing where movement is present near a keyboard in Fig 4.8, displaying correct results from our KNN model, and Fig 4.9 displays the correct classification of typing inside a difficult case where the monitor is partially occluding the keyboard.

### 4.2.3 Results for Talking

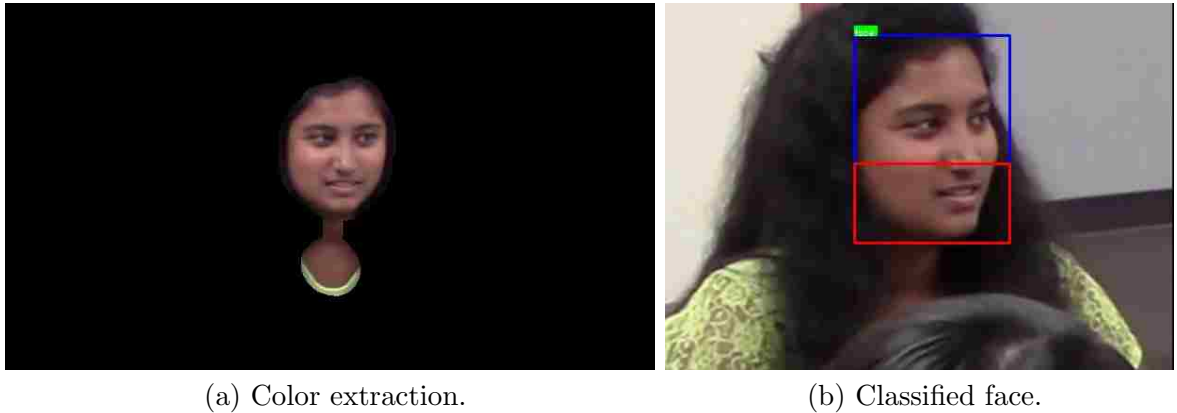


Figure 4.10: Subfigures and display correct classification results for the color and correct KNN results for the face.

In Fig 4.10 we see the two classification precursor steps for identifying talking successfully retrieved. First is the skin region necessary for finding the face in the KNN 4.10a, secondly we see the output of the KNN model during training after the face has been correctly identified in 4.10b.

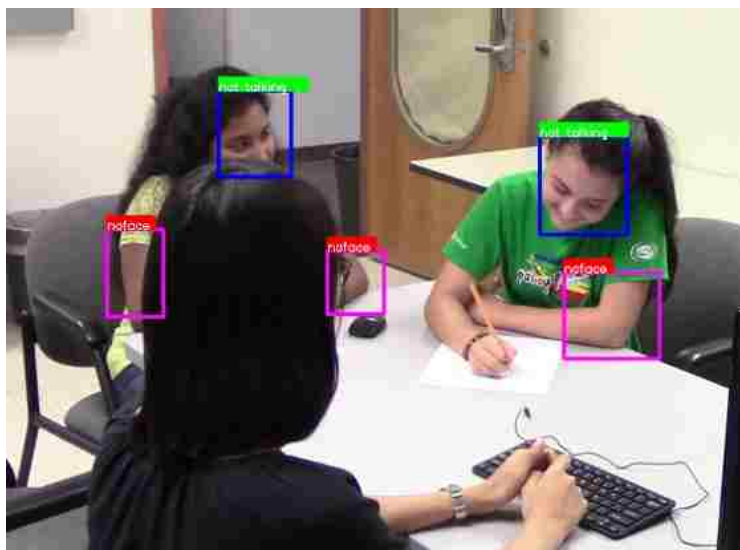
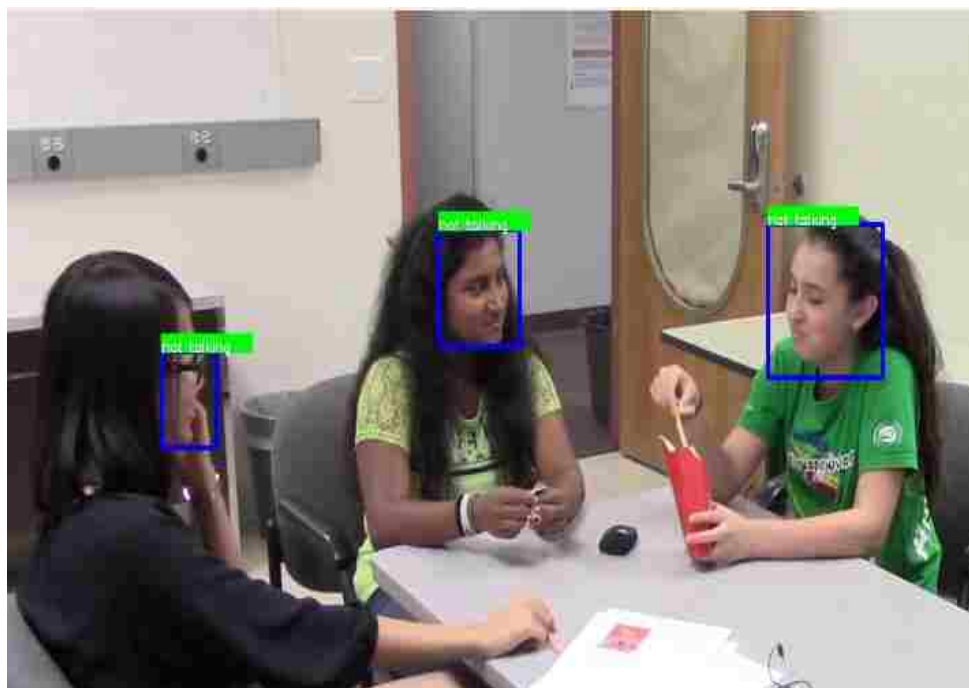


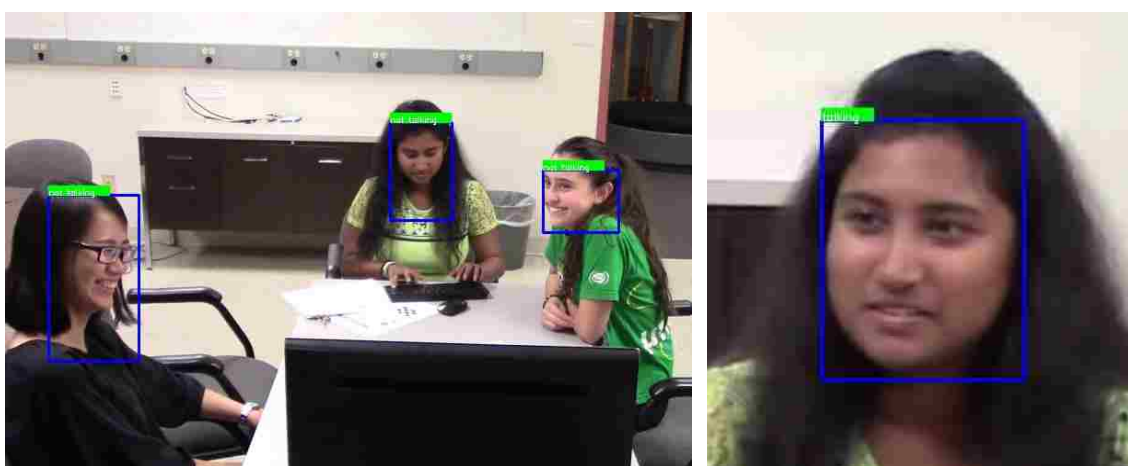
Figure 4.11: Correct classification of face KNN during validation.

Chapter 4. Results

Fig 4.11 displays correct face locations versus other skin areas being identified during the validation phase.



(a) No Talking.



(b) No Talking.

(c) Talking.

Figure 4.12: Subfigures and display correct classification results for no talking and talking.

## Chapter 4. Results

We present results of our validation set where talking has been successfully identified in three difficult cases. These cases are considered difficult due to the chances of laughing or eating being confused with talking.

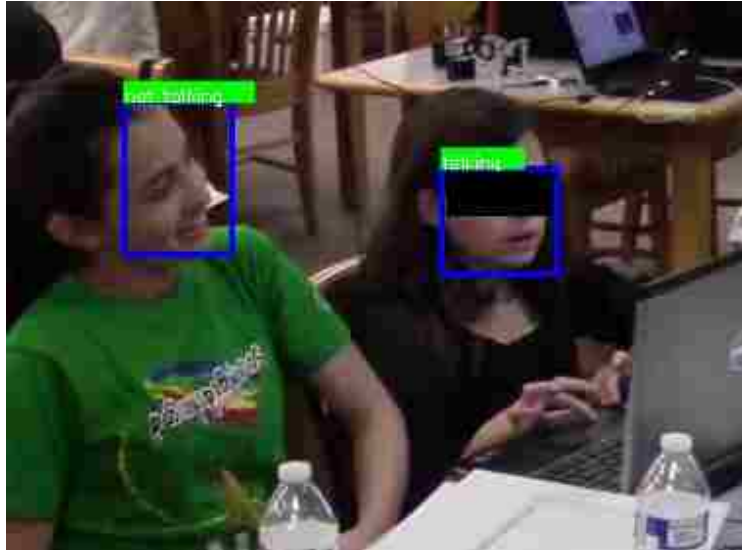


Figure 4.13: Correct classification of talking activity in difficult case in unseen video.

Finally, we present a difficult classification in unseen raw video from the AOLME dataset with accurate classification in Fig 4.13.

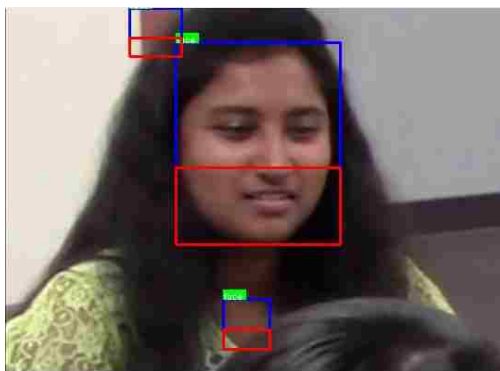
### 4.3 Failed Classification Examples

We present several failures of our models. First, we see a failure in the keyboard classification due to the color of the ledge of the table still being classified as a keyboard after passing through the KNN. The appearance of skin also confuses the model.

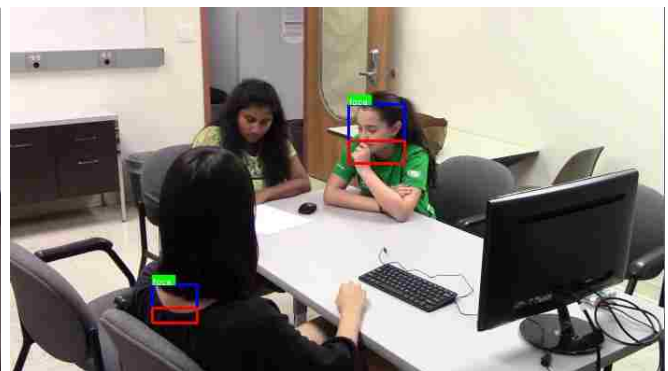
Chapter 4. Results



Figure 4.14: Incorrect classification of table and legs as keyboard.



(a) Face KNN Failure.



(b) Face KNN miss.

Figure 4.15: Subfigures and display incorrect functionality of the face KNN from skin and blank space.





Figure 4.16: Classification of a mouse as a keyboard.

We present three more KNN failures in object detection in Fig 4.15 and Fig 4.16, where objects have been misclassified using their respective KNN models. The features of these models are reviewed with their respective accuracies in Table 4.3. Keyboard classification accuracy was roughly 84% and faces were approximately 90% accurate in classification, with most confusing coming from solid areas of skin in the case of faces, and mice being confused for keyboards in the keyboard model.

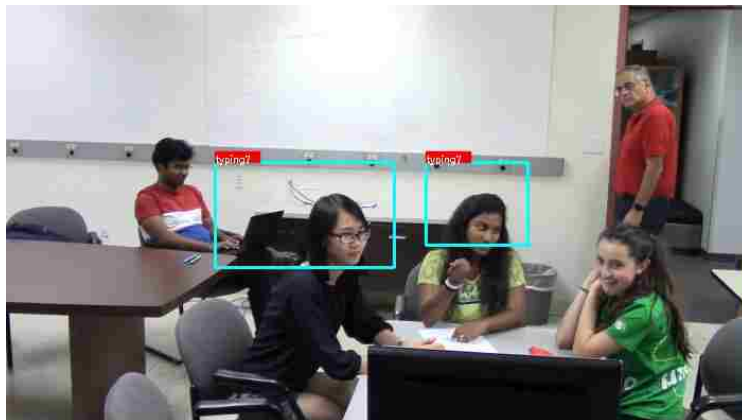


Figure 4.17: Table masking issue.



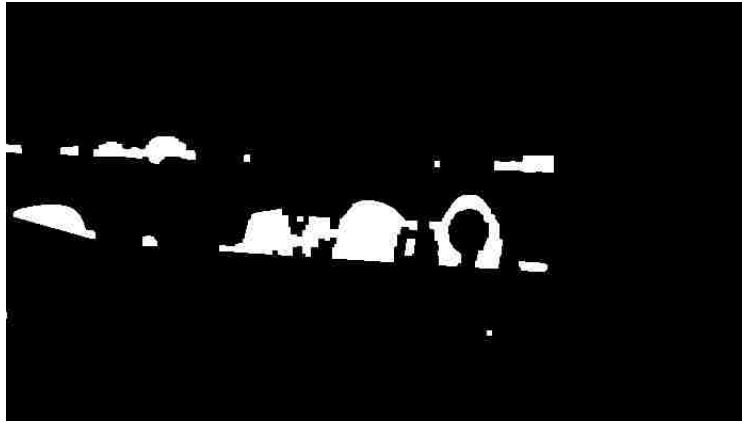


Figure 4.18: Incorrect masking due to whiteboard.

We present a total failure of object recognition in the two Figures 4.17 and 4.18, where the masking of the table has been confused by the whiteboard, and the classifier was passed information that was not actually on a table. This confused the classifier, which has reported everything as a keyboard.



Figure 4.19: Keyboard classified as mouse.

Additionally, we see another failure in training our KNN where the keyboard is seen to be a mouse in Fig 4.19. As can be seen in the results Table 4.3 containing the keyboard confusion matrix, this makes sense as most of our mice were classified

## Chapter 4. Results

as being keyboards, indicating we need a better training set for this data.

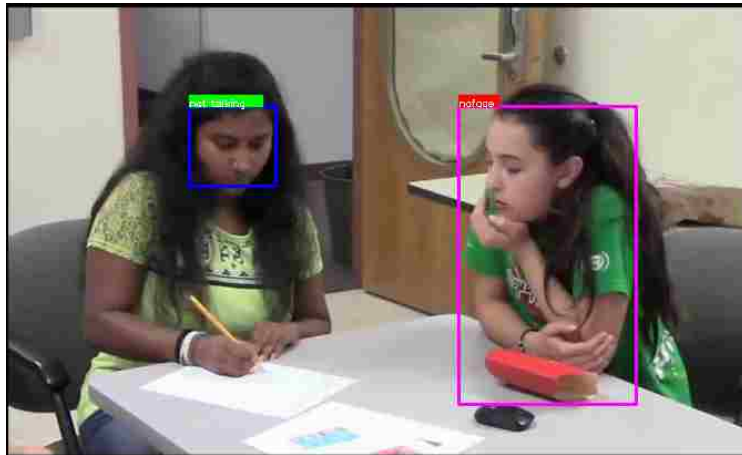


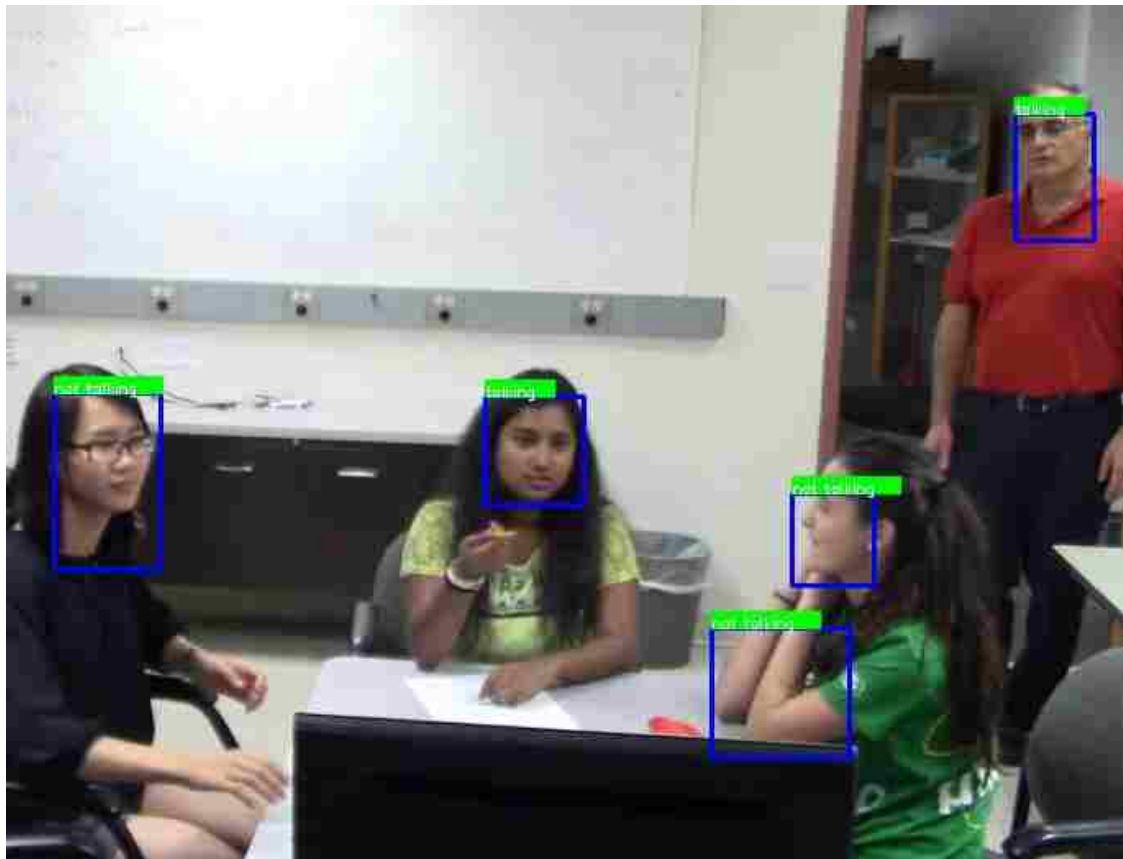
Figure 4.20: Face missed due to hand on face.

In Fig 4.20, we observe a classification that could not be made with the KNN model due to a hand being on the chin of the girl on the right. This kind of failure presents repeatedly in the research, especially in the unseen data. Having hands too close to faces, or someone else's arm too close to a face will cause a classification error as it confuses the KNN model, and additionally violates our area requirement stated in the previous section.

Chapter 4. Results



(a) Writing Failure. (b) Typing Failure



(c) Talking Failure.

Figure 4.21: Subfigures displaying incorrect classifications from all models.

Chapter 4. Results

We present failures of classification from all three models, which failed in the KNN and DNN methods. These were not difficult cases but likely had their motion confused due to a poor training set or otherwise.

	<b>Faces</b>	<b>Keyboards</b>
No of Positives	1209	75
No. of Negatives	399	25 Monitors, 21 Mice
Accuracy	89.75%	84.0%

Table 4.3: KNN Models for Typing and Talking

	Face	No Face
Face	233	15
No Face	18	56

Faces

	Key	Mon	Mou
Key	18	0	1
Mon	0	2	0
Mou	3	0	1

Keyboards

Figure 4.22: Confusion matrices for object KNN. Mon stands for Monitor.

	<b>Talking</b>	<b>Typing</b>	<b>Writing</b>
Accuracy	84.62%	70.96%	72.58%
K for KNN	11	18	15

Table 4.4: KNN Results

	<b>Talking</b>	<b>Typing</b>	<b>Writing</b>
Accuracy	52.99%	63.38%	70.16%
K for KNN	11	18	16

Table 4.5: Tensor-flow KNN Results

## Chapter 4. Results

	T	F		T	F		T	F
T	655	49	T	108	162	T	153	74
F	153	547	F	84	486	F	64	102
Talking			Typing			Writing		

Figure 4.23: Confusion matrices for KNN Results.

	<b>Talking</b>	<b>Typing</b>	<b>Writing</b>
Accuracy	99.72%	82.52%	92.47%
Batch Size	100	100	200
Activation Model	Relu	Selu	Selu
Neurons	100	50	70
Learning Rate	0.01	0.05	0.05
Hidden Layers	5	5	5
Regularization	L1/L2 Max_Norm	L1/L2 Max_Norm	L1/L2 Max_Norm

Table 4.6: DNN Results

## 4.4 Discussion

The classification results were found to be surprisingly accurate in some models and not as accurate in others. In the case of writing, we decided to stop and not further attempt to optimize the results due in part to a lack of time, but also because we wished to have a comparison to the models which included the KNN in addition to color to identify the other objects of interest. This gave us a great point of comparison to determine if the color of an object is enough to correctly identify a human activity.

In general, we found that color alone will work well under certain, inordinately specific circumstances, which include videos that are designed not unlike those in the most commonly used datasets. In these cases, we realize accurate classification because there is not much color noise in the background which could be misclassified

## *Chapter 4. Results*

as a pencil. The tabulated activity itself is found to be highly variable upon which videos are used to train the motion vectors. As many people have many different writing styles, some writing cannot be accurately classified using the motion vectors because the pencil will move so little. We see examples above of successful, good, and failure in the pencil classifier. The DNN model achieved the greatest results in terms of accuracy; however, often writing was still missed in certain cases, this is preferable to non-writing being inappropriately classified as writing. This activity detection model performed more poorly than the other two models.

Typing identification was moderately successful with most issues found in the classification of the keyboard itself, not unlike the major issues with the pencil model. Since the keyboard appears as a hole in the table, we had to take the convex hull of the table shape in an attempt to cover it to create context for the keyboard on the table. This did not always work due to the position of the monitor. If the monitor occluded any part of the keyboard, we were not able to ascertain whether a keyboard was present or not. Additionally, no classification takes place in the situation where the hull does not cover the entirety of the table region, which occurs when the monitor blocks both edges of the table. We found often our keyboard was misclassified as a mouse, and this caused several misclassifications during the training phase. Again, this was highly variable upon the dataset used for training due to different typing styles those who use the whole hand to type versus hunt and peck and whether the mouse was placed too closely to the keyboard. This model was found to be more accurate than the writing model; so long as the keyboard is identified, the typing is correctly labelled. Again, the DNN outperforms both KNN models.

For talking, we had the fewest errors. The faces were far easier to identify than the keyboard and thus finding the motion of the bottom half of the face was not difficult. Due to this, we get an incredibly high accuracy on all our models. There are some specific problems when the face is not correctly identified, usually as an

## *Chapter 4. Results*

arm or a hand, which are most often classified as not talking, so not much change is visually observed when reviewing playback unless the debug labels are applied. This model performed best when applied to unseen datasets that were not used in training or testing. The DNN performed exceedingly well in this model, though the KNN accuracies are also quite close.

The videos used in training and testing varied between experiments, the majority of training being done on a collection of videos specifically recorded for the purposes of the ECE 633 Advanced Image Processing class at UNM. Some video from the AOLME dataset were also used in training the classifiers, although we attempted to reserve as many of these as possible for the purpose of testing unseen video data.

For the writing videos, a cropped dataset were used where each video was cut so the pencil would be clearly seen in each frame containing writing and raw video was used for training no writing. This was done so the writing videos would only record the motions of the pencil and the training set would contain no noise.

For the typing dataset, the full uncut videos were used for training of typing and no typing, though were specifically selected so that the keyboard was not obscured by the monitor. This was not realistic in the real-world video thus caused a lot of classification error when applied to the raw dataset.

For talking, closeups of people talking were used to record motion vectors of the mouth moving, again uncut video was used to record faces that were not speaking, but moving around.

When each model is applied to raw datasets we achieved some amount of success, the greatest of which being seen with talking versus not talking. Since we obtained extremely high accuracy rates with the talking vs. not talking, this was a somewhat expected result. It is believed this is due to how people tend to move their heads when speaking, and hold mostly still while doing anything else, except in the case

## *Chapter 4. Results*

of reading. The other videos seemed to have a variable amount of error with each model, most of which lasted only a few frames before self-correcting.

Finally, the datasets used for testing and training came from different video databases, and were capable of correctly identifying the specified motions in different settings with different actions without editing. This also was somewhat successful when used on videos obtained from imagenet for writing, which means the model could potentially generalize with some amount of optimization.



# Chapter 5

## Conclusions and Future Work

### 5.1 Conclusions

We conclude that by combining object context with motion vector training, we can classify human activity within a set of unrelated raw videos with some degree of accuracy. The method can locate an actions specific location in a busy video in which many other activities are taking place. There were a few obstacles to increasing the accuracy in our model which may easily be overcome with an applied amount of optimization and a far larger and more diverse video training set.

The largest errors to overcome in our model concern appropriate identification of objects we wish to apply context to before extracting the motion vectors. These include errors present when objects are the same color or tone as the primary object we are attempting to locate in a frame. When there is another object that is similar color resembling one of the interaction objects it confuses the classifier, and simple movements like a head passing by will cause inaccurate identification of writing. We attempted to correct this with the keyboard and face by using an additional KNN to verify we are adding context to the correct object. The extra context does improve

## *Chapter 5. Conclusions and Future Work*

accuracy, but even with the extra processing it still is not perfect.

The monitor created several problems for us in classification videos as it tended to block line of sight to many activities we were seeking. This, unfortunately, does not have a solution as we cannot see through objects. Most of the problems were seen with typing, since the monitor was often in the way of the keyboard and they are the same color, so they often blended together when the colors were extracted for thresholding, therefore we could not accurately locate the keyboard in the frame in all cases.

Though we obtained a  $>99.6\%$  classification accuracy using a DNN trained for talking, we do not trust this number as it seems far too high accuracy of a classification rate. We believe this deserves extreme scrutiny in any future study, perhaps removal of dense vectors which are associated with head movement. The classification accuracy averages in the mid 70% for writing and typing activity using KNN seem far more realistic, and when played back on a raw dataset unrelated to our training videos, seems to match the amount of errors per frame we see during playback. We chose to use the average accuracy in these cases rather than the maximum accuracies obtained since they were not reproducible after adding to the training set.

Overall, the method is a promising step toward accurate classification of human activity as it eliminates much of the confusion of action that is seen in current state-of-the-art models by giving a location and context to the objects in a frame. The generality of the model allows it to be used on more than one dataset, and with some additional work could potentially serve as the first step in automatic classification procedures in the future.

Furthermore, we conclude that KNN is not the best classification method for this type of dataset. Due to the nature of the motions being quite minute in magnitude, many times motion vectors for a minute action are confused with noise or other

movements, which lead to misclassification of illumination noise as movement, an unfortunate side effect of the way dense optical flow operates.

## **5.2 Future Work**

Given more time, we would like to apply a RBF neural network to classify the found object motion, eliminating the need for cropping video and labelling the motion vectors. Since RBFs can come up with a classification without this data, it would facilitate a more robust model. As KNN model performed poorly, a better model for classification may be a SVM or a Clustering based algorithm with some amount of dimensionality reduction, and the RBF would be an ideal replacement given those conditions.

Many issues in our method came from having too much color noise in the frame being classified as an action when it was too close to other similar objects. Though we attempted to apply a median blur to each frame to remove some illumination noise, it could not all be reduced, causing many classification errors. Lucas-Kanade should be investigated as an alternative to dense optical flow.

Another interesting addition would be to add haar cascades for object recognition as this would allow for classification of more than a single type of pencil, or single type of mouse, etc. Our model was highly limited by the fact that we only sought objects of the same color and having a library of objects to look at could greatly improve the generalization of the model in the future.

Finally, the application of an AI such as PAL at the end of the DNN classifier, we could insert a random set of objects and allow the AI to search for the objects that are found to be interacting with one another, thereby giving it a more appropriate label. For example, hands sideways pencil could be classified as drawing, or hands,

## *Chapter 5. Conclusions and Future Work*

keyboard, table can be found as typing. This effectively eliminates the need for writing complex context functions and the necessitation of labelling video data, since when linked with an RBF for unsupervised learning, could automatically label the interactions and activities. This would also help with identifying noisy regions so that they could more easily be ignored in any future model.

# Chapter 6

## Appendices

# Appendix A

## Python Code for Color Detection

### A.1 Paper

*#Paper portion of code*

```
def FindPaper(hsv, frame, an_object):
```

```
    # Method provides maximum customization options
```

```
    kernel = cv2.getStructuringElement(cv2.MORPH_ELLIPSE,
```

```
    #ideally these will be defined by a file that has these values stored
```

```
    #or more ideally a file will define all the code.
```

```
    mask1 = cv2.inRange(hsv, (0, 0, 160), (188, 20, 240))
```

```
    mask1 = cv2.erode(mask1, None, iterations=2)
```

```
    mask1 = cv2.dilate(mask1, None, iterations=2)
```

```
    mask = cv2.inRange(hsv, (0, 0, 174), (180, 255, 255))
```

```
    mask = cv2.erode(mask, None, iterations=2)
```

```
    mask = cv2.dilate
```

```
    res1 = cv2.bitwise_and(frame, frame, mask= mask1)
```

## Appendix A. Python Code for Color Detection

```
res2 = cv2.bitwise_and(frame, frame, mask= mask)
res2 = res2 - res1
greyfr2 = cv2.cvtColor(res2, cv2.COLOR_BGR2GRAY)
_, thr2 = cv2.threshold(greyfr2, 210, 255,
                        cv2.THRESH_BINARY)
thr2 = cv2.morphologyEx(thr2, cv2.MORPH_CLOSE,
                        kernel, iterations=2)
thr2 = cv2.morphologyEx(thr2, cv2.MORPH_OPEN, kernel)
return thr2
```

## A.2 Pencil

*#Pencil portion of code*

```
def FindPencil(hsv, frame, an_object):
    t = cv2.cvtColor(frame, cv2.COLOR_BGR2YCR_CB)
    mask2 = cv2.inRange(hsv, (13, 45, 60), (30, 176, 255))
    mask2 = cv2.erode(mask2, None, iterations=1)
    mask2 = cv2.dilate(mask2, kernel, iterations=3)
    mask2 = cv2.morphologyEx(mask2, cv2.MORPH_CLOSE,
                             kernel, iterations=4)
    mask2 = cv2.GaussianBlur(mask2, (11, 11), 0)
    mask3 = cv2.inRange(t, (69, 125, 0), (254, 199, 105))
    mask3 = cv2.erode(mask3, None, iterations=1)
    mask3 = cv2.dilate(mask3, kernel, iterations=3)
    mask3 = cv2.morphologyEx(mask3, cv2.MORPH_CLOSE,
                             kernel, iterations=4)
    mask3 = cv2.GaussianBlur(mask3, (11, 11), 0)
    res = cv2.bitwise_and(frame, frame, mask= mask3)
```

## Appendix A. Python Code for Color Detection

```
greyfr = cv2.cvtColor(res ,cv2.COLOR_BGR2GRAY)
_,thr = cv2.threshold(greyfr ,130 ,240 ,
                      cv2.THRESH_BINARY)

th2 = PersonObject(frame)
th2 = cv2.GaussianBlur(th2 , (11 , 11) , 0)
th2 = cv2.dilate(th2 , kernel , iterations=3)
gray = cv2.cvtColor(th2 ,cv2.COLOR_BGR2GRAY)
_,th2 = cv2.threshold(gray , 0 ,255 ,
                      cv2.THRESH_BINARY)

return thr
```

### A.3 Keyboard

*#Keyboard portion of code*

```
def FindKeyboard(hsv ,frame ,an_object ):
    min_YCrCb = np.array ([0 , 120 , 127] ,np.uint8)
    max_YCrCb = np.array ([144 , 255 , 177] ,np.uint8)
    # Convert image to YCrCb
    image_YCrCb = cv2.cvtColor(frame ,
                                cv2.COLOR_BGR2YCR_CB)
    dark_region = cv2.inRange(image_YCrCb ,
                               min_YCrCb ,max_YCrCb)
    res3 = cv2.erode(dark_region , None , iterations=2)
    res3 = cv2.dilate(res3 , None , iterations=2)
    res3 = cv2.bitwise_and(frame ,frame , mask = res3)
    greyfr3 = cv2.cvtColor(res3 ,cv2.COLOR_BGR2GRAY)
    _,thr3 = cv2.threshold(greyfr3 ,0 ,255 ,
                           cv2.THRESH_BINARY)
```



## Appendix A. Python Code for Color Detection

```
return thr3, greyfr3
```

### A.4 Table

*#Table portion of code*

```
def FindTable(hsv, frame, an_object):
    min_YCrCb = np.array([181, 109, 127], np.uint8)
    max_YCrCb = np.array([255, 135, 160], np.uint8)
    # Convert image to YCrCb
    image_YCrCb = cv2.cvtColor(frame,
                                cv2.COLOR_BGR2YCR_CB)
    mask = cv2.inRange(image_YCrCb, min_YCrCb, max_YCrCb)
    res2 = cv2.bitwise_and(frame, frame, mask= mask)
    greyfr3 = cv2.cvtColor(res2, cv2.COLOR_BGR2GRAY)
    _, thr3 = cv2.threshold(greyfr3, 0, 255,
                            cv2.THRESH_BINARY)
    thr3 = cv2.morphologyEx(thr3, cv2.MORPH_CLOSE,
                            kernel, iterations = 3)
    thr3 = cv2.morphologyEx(thr3, cv2.MORPH_OPEN,
                            kernel, iterations = 2)
    return thr3
```

### A.5 Skin

```
def PersonObject(frame):
    kernel = cv2.getStructuringElement(cv2.MORPH_ELLIPSE,
                                       (11, 11))
    f = np.copy(frame)
```

## Appendix A. Python Code for Color Detection

```
min_YCrCb = np.array([75,138,111],np.uint8)
max_YCrCb = np.array([172,155,124],np.uint8)
min_YCrCb1 = np.array([56, 142, 110],np.uint8)
image_YCrCb = cv2.cvtColor(f,cv2.COLOR_BGR2YCrCb)
imgYCC = cv2.GaussianBlur(image_YCrCb, (11, 11), 0)
# Find region with skin tone in YCrCb image
skin_region = cv2.inRange(imgYCC,min_YCrCb,max_YCrCb)
skin_region2 = cv2.inRange(imgYCC,min_YCrCb1,max_YCrCb)
skin_region = cv2.bitwise_and(skin_region,skin_region,
                              mask = skin_region2)
skin_region = cv2.morphologyEx(skin_region,
                              cv2.MORPHOPEN, kernel)
skin_region = cv2.dilate(skin_region, None,
                        iterations=10)
skin_region = cv2.morphologyEx(skin_region,
                              cv2.MORPHCLOSE, kernel)
skin_region = cv2.GaussianBlur(skin_region, (29,29), 0)
skin_region = cv2.erode(skin_region, kernel2,
                        iterations=2)
skin_region = cv2.morphologyEx(skin_region,
                              cv2.MORPHOPEN, kernel,
                              iterations=2)
res = cv2.bitwise_and(f,f,mask=skin_region)
return res
```

# References

- [1] *BogoToBogo object detection : Face detection using haar cascade classifiers*. Accessed: 2017-10-07.
- [2] *DevPost human motion classification*. Accessed: 2017-09-02.
- [3] *GitHub artificial-intelligence-deep-learning-machine-learning-tutorials*. Accessed: 2017-10-15.
- [4] *GitHub tensorflow\_tutorials*. Accessed: 2017-10-15.
- [5] *Machine Learning Mastery handwritten digit recognition using convolutional neural networks in python with keras*. Accessed: 2017-10-18.
- [6] *Medium building a real-time object recognition app with tensorflow and opencv*. Accessed: 2017-10-18.
- [7] *MEMEMEME training haar cascades*. Accessed: 2017-08-05.
- [8] *Naotoshi Seo tutorial: opencv haartraining (rapid object detection with a cascade of boosted classifiers based on haar-like features)*. Accessed: 2017-10-05.
- [9] *PyImage Search*. Accessed: 2017-07-03.
- [10] *Tensor Flow api master*. Accessed: 2017-09-10.
- [11] S. Acton, P. Soliz, S. Russell, and M. Pattichis, *Content based image retrieval: The foundation for future case-based and evidence-based ophthalmology*, Proc. IEEE international conference on multimedia and expo, 2008, June, pp. 541–544.

## REFERENCES

- [12] C. Agurto, S. Barriga, V. Murray, S. Murillo, G. Zamora, W. Bauman, M. Pattichis, and P. Soliz, *Toward comprehensive detection of sight threatening retinal disease using a multiscale am-fm methodology*, Spie medical imaging, 2011, pp. 796316–796316.
- [13] C. Agurto, S. Barriga, V. Murray, S. Nemeth, R. Crammer, W. Bauman, G. Zamora, M. Pattichis, and P. Soliz, *Automatic detection of diabetic retinopathy and age-related macular degeneration in digital fundus images*, Investigative Ophthalmology and Visual Science **52** (2011), no. 8, 5862–5871.
- [14] C. Agurto, S. Barriga, V. Murray, M. Pattichis, and P. Soliz, *Effects of image compression and degradation on an automatic diabetic retinopathy screening algorithm*, Proc. spie medical imaging: Computer-aided diagnosis, 2010, pp. 76240H–76240H.
- [15] C. Agurto, S. Murillo, V. Murray, M. Pattichis, S. Russell, M. Abramoff, and P. Soliz, *Detection and phenotyping of retinal disease using am-fm processing for feature extraction*, 42nd asilomar conference on signals, systems and computers, 2008Oct, pp. 659–663.
- [16] C. Agurto, V. Murray, S. Barriga, S. Murillo, M. Pattichis, H. Davis, S. Russell, M. Abramoff, and P. Soliz, *Multiscale am-fm methods for diabetic retinopathy lesion detection*, IEEE Transactions on Medical Imaging **29** (2010Feb), no. 2, 502–512.
- [17] C. Agurto, V. Murray, H. Yu, J. Wigdahl, M. Pattichis, S. Nemeth, S. Barriga, and P. Soliz, *A multiscale optimization approach to detect exudates in the macula*, IEEE Journal of Biomedical and Health Informatics **18** (2014July), no. 4, 1328–1336.
- [18] C. Agurto, H. Yu, V. Murray, M. Pattichis, S. Barriga, W. Bauman, and P. Soliz, *Detection of neovascularization in the optic disc using an am-fm representation, granulometry, and vessel segmentation*, Ieee embc, 2012Aug, pp. 4946–4949.
- [19] C. Agurto, H. Yu, V. Murray, M. Pattichis, S. Barriga, and P. Soliz, *Detection of hard exudates and red lesions in the macula using a multiscale approach*, Ieee southwest symposium on image analysis and interpretation (ssiai), 2012April, pp. 13–16.

## REFERENCES

- [20] C. Agurto, H. Yu, V. Murray, M. Pattichis, S. Nemeth, S. Barriga, and P. Soliz, *A multiscale decomposition approach to detect abnormal vasculature in the optic disc*, Computer Graphics and Image Processing **42** (2015July), 137–149.
- [21] Z. Antoniou, A. Panayides, M. Pantziaris, A. Constantinides, C. Pattichis, and Pattichis M., *Dynamic network adaptation for real-time medical video communication*, Xiv mediterranean conference on medical and biological engineering and computing, 2016.
- [22] Z. Antoniou, A. Panayides, and M.S. Pattichis, *Adaptive emergency scenery video communications using hevc for responsive decision support in disaster incidents*, 37th annual international conference of the ieee engineering in medicine and biology society (embc), 2015, pp. 173–176.
- [23] M. Arbabshirani, M. Pattichis, A. Ulloa, and V. Calhoun, *Detecting volumetric changes in fmri connectivity networks in schizophrenia patients*, 36th ieee embc, 2014Aug, pp. 726–729.
- [24] S. Barriga, V. Murray, C. Agurto, M. Pattichis, W. Bauman, G. Zamora, and P. Soliz, *Automatic system for diabetic retinopathy screening based on am-fm, partial least squares, and support vector machines*, Ieee international symposium on biomedical imaging, 2010April, pp. 1349–1352.
- [25] S. Barriga, V. Murray, C. Agurto, M. Pattichis, S. Russell, M. Abramoff, H. Davis, and P. Soliz, *Multi-scale am-fm for lesion phenotyping on age-related macular degeneration*, 22nd ieee symposium on computer-based medical systems, 2009Aug, pp. 1–5.
- [26] S. Barriga, M. Pattichis, M. Abramoff, D. Ts'o, R. Kardon, Y. Kwon, and P. Soliz, *Spatiotemporal independent component analysis for retinal images*, 40th asilomar conference on signals, systems and computers, 2006Oct, pp. 1961–1965.

## REFERENCES

- [27] ———, *Spatiotemporal independent component analysis for the detection of functional responses in cat retinal images*, IEEE Transactions on Medical Imaging **26** (2007Aug), no. 8, 1035–1045.
- [28] S. Barriga, M. Pattichis, M. Abramoff, D. Ts'o, Y. Kwon, R. Kardon, and P. Soliz, *Independent component analysis for the detection of in-vivo intrinsic signals from an optical imager of retinal function*, Proc. of photonics west, 2007, pp. 64261K–64261K.
- [29] S. Barriga, M. Pattichis, D. Ts'o, M. Abramoff, Randy Kardon, Young Kwon, and P. Soliz, *Independent component analysis using prior information for signal detection in a functional imaging system of the retina*, Medical image analysis **15** (2011), no. 1, 35–44.
- [30] S. Barriga, P. Truitt, M. Pattichis, D. Ts'o, Y. Kwon, R. Kardon, and P. Soliz, *Blind source separation in retinal videos*, Proc. of the spie medical imaging conference, 2003, pp. 1591–1601.
- [31] S. Barriga, D. Ts'o, M. Pattichis, Y. Kwon, R. Kardon, M. Abramoff, and P. Soliz, *Detection of low-amplitude in vivo intrinsic signals from an optical imager of retinal function*, Proc. spie photonics, 2006, pp. 61380F–61380F–12.
- [32] S. Barriga, D. Ts'o, M. Pattichis, and P. Soliz, *Independent component analysis for processing of retinal responses to patterned stimuli*, Proc. of the 25th international conference of the ieee engineering in medicine and biology society, 2003Sept, pp. 1006–1009 Vol.1.
- [33] T. Cacoullou and M. Pattichis, *Classification under a multivariate bernoulli: An application to pneumoconiosis*, Proc. of the panhellenic conference on statistics, 2005, pp. 419–426.
- [34] ———, *Classification under a multivariate bernoulli: An application to pneumoconiosis*, Proc. of spie medical imaging, 2006.
- [35] ———, *Pneumoconiosis revisited: Classifier viewed via roc curves and logic functions*, Proc. of the panhellenic conference on statistics, 2006.

## REFERENCES

- [36] C. Carranza, D. Llamocca, and M. Pattichis, *A scalable architecture for implementing the fast discrete periodic radon transform for prime sized images*, Ieee international conference on image processing, 2014Oct, pp. 1208–1212.
- [37] ———, *Fast and scalable computation of the forward and inverse discrete periodic radon transform*, IEEE Transactions on Image Processing **25** (2016Jan), no. 1, 119–133.
- [38] C. Carranza, D. Llamocca, and M.S. Pattichis, *The fast discrete periodic radon transform for prime sized images: Algorithm, architecture, and vlsi/fpga implementation*, Ieee southwest symposium on image analysis and interpretation, 2014april, pp. 169–172.
- [39] C. Carranza, V. Murray, M. Pattichis, and S. Barriga, *Multiscale am-fm decompositions with gpu acceleration for diabetic retinopathy screening*, Ieee southwest symposium on image analysis and interpretation (ssiai), 2012April, pp. 121–124.
- [40] Cesar Carranza, Daniel Llamocca, and Marios Pattichis, *Fast 2d convolutions and cross-correlations using scalable architectures*, IEEE Transactions on Image Processing **26** (2017), no. 5, 2230–2245.
- [41] S. Celedón-Pattichis, C. LópezLeiva, M. Pattichis, and D. Llamocca, *An interdisciplinary collaboration between computer engineering and mathematics/bilingual education to develop a curriculum for underrepresented middle school students*, Cultural Studies of Science Education **8** (2013), no. 4, 873–887.
- [42] C. Christodoulou, P. Kaplanis, V. Murray, M. Pattichis, and C. Pattichis, *Classification of surface electromyographic signals using am-fm features*, 9th international conference on information technology and applications in biomedicine, 2009Nov, pp. 1–4.
- [43] ———, *Comparison of am-fm features with standard features for the classification of surface electromyographic signals*, Xii mediterranean conference on medical and biological engineering and computing, 2010, pp. 69–72.

## REFERENCES

- [44] C. Christodoulou, P. Kaplanis, V. Murray, M. Pattichis, C. Pattichis, and T. Kyriakides, *Multi-scale amfm analysis for the classification of surface electromyographic signals*, Biomedical Signal Processing and Control **7** (2012), no. 3, 265–269.
- [45] C. Christodoulou, E. Kyriacou, M. Pattichis, and C. Pattichis, *Plaque feature extraction*, Ultrasound and carotid bifurcation atherosclerosis, 2012, pp. 223–246.
- [46] C. Christodoulou, E. Kyriacou, M. Pattichis, C. Pattichis, and A. Nicolaidis, *A comparative study of morphological and other texture features for the characterization of atherosclerotic carotid plaques*, Proc. of the 10th international conference on computer analysis of images and patterns, 2003, pp. 503–511 (English).
- [47] ———, *A comparative study of morphological and other texture features for the characterization of atherosclerotic carotid plaques*, Computer analysis of images and patterns, lecture notes in computer science (lncs2756), 2003, pp. 503–511.
- [48] C. Christodoulou, C. Loizou, C. Pattichis, M. Pantziaris, E. Kyriacou, M. Pattichis, C. Schizas, and A. Nicolaidis, *De-speckle filtering in ultrasound imaging of the carotid artery*, Ieee 24th international conference of the ieee engineering in medicine and biology society, 2002, pp. 1027–1028 vol.2.
- [49] C. Christodoulou, C. Pattichis, E. Kyriacou, M. Pattichis, M. Pantziaris, and A. Nicolaidis, *Ultrasound imaging in the assessment of carotid plaque morphology*, Medical image analysis methods, 2005, pp. 87–135.
- [50] C. Christodoulou, C. Pattichis, V. Murray, M. Pattichis, and A. Nicolaidis, *Am-fm representations for the characterization of carotid plaque ultrasound images*, Proc. 4th european conference of the international federation for medical and biological engineering, 2009, pp. 546–549.
- [51] I. Constantinou, M. Neofytou, V. Tanos, M. Pattichis, C. Christodoulou, and C. Pattichis, *A comparison of color correction algorithms for endoscopic cameras*, 13th ieee international conference on bioinformatics and bioengineering, 2013Nov, pp. 1–4.



## REFERENCES

- [52] I. Constantinou, M. Pattichis, and C. Pattichis, *Multiscale am-fm image reconstructions based on elastic net regression and gabor filterbanks*, Asilomar conference on signals, systems and computers, 2013Nov, pp. 1985–1989.
- [53] I. Constantinou, M. Pattichis, V. Tanos, M. Neofytou, and C. Pattichis, *An adaptive multiscale am-fm texture analysis system with application to hysteroscopy imaging*, Ieee 12th international conference on bioinformatics bioengineering (bibe), 2012Nov, pp. 744–747.
- [54] I. Constantinou, M. Pattichis, C. Tziakouri, C. Pattichis, and S. Petroudi, *Multiscale am-fm models and instantaneous amplitude evaluation for mammographic density classification*, Medical image understanding and analysis (miua), 2014.
- [55] I. Constantinou, M. Pattichis, C. Tziakouri, C. Pattichis, S. Petroudi, and C. Nicosia, *Multiscale am-fm models and instantaneous amplitude evaluation for mammographic density classification.*, Miua, 2014, pp. 271–276.
- [56] Jeff Donahue, Lisa Anne Hendricks, Sergio Guadarrama, Marcus Rohrbach, Subhashini Venugopalan, Kate Saenko, and Trevor Darrell, *Long-term recurrent convolutional networks for visual recognition and description*, CoRR **abs/1411.4389** (2014), available at [1411.4389](https://arxiv.org/abs/1411.4389).
- [57] Cody W Eilar, Venkatesh Jatla, Marios S Pattichis, Carlos LópezLeiva, and Sylvia Celedón-Pattichis, *Distributed video analysis for the advancing out of school learning in mathematics and engineering project*, 50th asilomar conference on signals, systems and computers, 2016, pp. 571–575.
- [58] Cody Wilson Eliar, *Distributed and scalable video analysis architecture for human activity recognition using cloud services*, Master’s Thesis, 2016.
- [59] G. Esakki, V. Jatla, and M. Pattichis, *Optimal hevcc encoding based on gop configurations*, Ieee southwest symposium on image analysis and interpretation, 2016.
- [60] Aurelien Geron, *Hands-on machine learning with scikit-learn and tensorflow*, OReilly, 2017.

## REFERENCES

- [61] T. Gevers and A. W. M. Smeulders, *Color based object recognition* (Alberto Del Bimbo, ed.), Springer Berlin Heidelberg, Berlin, Heidelberg, 1997.
- [62] J. Havlicek, M. Pattichis, D. Harding, A. Christofides, and A. Bovik, *Am-fm image analysis techniques*, Proc. of the ieee southwest symposium on image analysis and interpretation, 1996Apr, pp. 195–200.
- [63] V.M.M. Herrera, M.S. Pattichis, P. Soliz, C.P.A. Rios, and H.T. Davis, *System and methods of amplitude-modulation frequency-modulation (am-fm) demodulation for image and video processing*, Google Patents, 2013. US Patent 8,515,201.
- [64] J. Hoffman and M. Pattichis, *A high-speed dynamic partial reconfiguration controller using direct memory access through a multiport memory controller and overclocking with active feedback*, International Journal of Reconfigurable Computing **2011** (2011).
- [65] Y. Honggang, M. Pattichis, C. Agurto, and B. Goens, *A 3d freehand ultrasound system for multi-view reconstructions from sparse 2d scanning planes*, Biomedical engineering online **10** (2011), no. 1, 7.
- [66] A. Jakomulska and K. Clarke, *Variogram-derived measures of textural image classification*, geoenv iii – geostatistics for environmental applications, 2001, pp. 345–355 (English).
- [67] O. Jeromin, V. Calhoun, and M. Pattichis, *Optimal sampling geometries for tv-norm reconstruction of fmri data*, 42nd asilomar conference on signals, systems and computers, 2008Oct, pp. 1397–1401.
- [68] O. Jeromin and M. Pattichis, *Reconstruction of aerial image from fourier spectral samples using statistical models*, Proc. 43rd asilomar conference on signals, systems and computers, 2009, pp. 1397–1401.
- [69] ———, *Multiscale sampling geometries and methods for deterministic and stochastic reconstructions of magnitude and phase spectra of satellite imagery*, IEEE Transactions on Geoscience and Remote Sensing **50** (2012), no. 10, 3678–3692.

## REFERENCES

- [70] O. Jeromin, M. Pattichis, and V. Calhoun, *Optimal compressed sensing reconstructions of fmri using 2d deterministic and stochastic sampling geometries*, Biomedical engineering online **11** (2012), no. 1, 25.
- [71] Y. Jiang, G. Esakki, and M. Pattichis, *Dynamically reconfigurable architecture system for time-varying image constraints (drastic) for hevc intra encoding*, Asilomar conference on signals, systems and computers, 2013Nov, pp. 1112–1116.
- [72] Y. Jiang, D. Llamocca, M. Pattichis, and G. Esakki, *A unified and pipelined hardware architecture for implementing intra prediction in hevc*, Ieee southwest symposium on image analysis and interpretation, 2014April, pp. 29–32.
- [73] Y. Jiang and M. Pattichis, *Jpeg image compression using quantization table optimization based on perceptual image quality assessment*, Proc. 45th asilomar conference on signals, systems and computers, 2011Nov, pp. 225–229.
- [74] ———, *A dynamically reconfigurable dct architecture for maximum image quality subject to dynamic power and bitrate constraints*, Ieee southwest symposium on image analysis and interpretation (ssiai), 2012April, pp. 189–192.
- [75] ———, *Dynamically reconfigurable dct architectures based on bitrate, power, and image quality considerations*, 19th ieee international conference on image processing (icip), 2012Sept, pp. 2465–2468.
- [76] ———, *A dynamically reconfigurable deblocking filter for h.264/avc codec*, Asilomar conference on signals, systems and computers, 2012Nov, pp. 2189–2193.
- [77] ———, *A dynamically reconfigurable architecture system for time-varying image constraints (drastic) for motion jpeg*, Journal of Real-Time Image Processing (2014), 1–17.
- [78] Y. Jiang, T. Perez, M. Pattichis, and B. Khan, *A configurable system for role-specific video imaging during laparoscopic surgery*, Ieee 12th international conference on bioinformatics bioengineering (bibe), 2012Nov, pp. 738–743.

## REFERENCES

- [79] Y. Jiang, C. Zong, and M.S. Pattichis, *Scalable hevc intra frame complexity control subject to quality and bitrate constraints*, 3rd ieee global conference on signal and information processing, 2015. In press.
- [80] R. Jordan and M. Pattichis, *Fast am-fm demodulation image and video analysis using single instruction multiple data (simd) architectures*, Proc. iasted international conference on signal processing, pattern recognition and applications, spr., 2003, pp. 230–233.
- [81] J. Kern and M. Pattichis, *Robust multispectral image registration using mutual-information models*, IEEE Transactions on Geoscience and Remote Sensing **45** (2007May), no. 5, 1494–1505.
- [82] J. Kern, M. Pattichis, and S. Stearns, *Registration of image cubes using multivariate mutual information*, Proc. 37th asilomar conference on signals, systems and computers, 2003Nov, pp. 1645–1649 Vol.2.
- [83] J. Kief, M. Pattichis, H. Pollard, A. Vera, and E. Parra, *An xup-unm educational platform: A dual-fpga platform for reconfigurable logic*, Computer Applications in Engineering Education **17** (2009), no. 2, 232–239.
- [84] E. Kyriacou, C. Christodoulou, C. Loizou, M. Pattichis, C. Pattichis, S. Kakkos, and A. Nicolaidis, *Assessment of stroke by analyzing carotid plaque morphology*, Handbook of research on advanced techniques in diagnostic imaging and biomedical applications, 2009, pp. 21.
- [85] E. Kyriacou, C. Christodoulou, M. Pattichis, C. Pattichis, and S. Kakkos, *Plaque classification*, Ultrasound and carotid bifurcation atherosclerosis, 2012, pp. 247–262.
- [86] E. Kyriacou, A. Constantinides, C. Pattichis, M. Pattichis, and A. Panayides, *eemergency health care information systems*, 33rd annual international conference of the ieee engineering in medicine and biology society, ieee embc, 2011Aug, pp. 2501–2504.
- [87] ———, *eemergency healthcare informatics*, Biomedical signals, imaging, and informatics, 2015.

## REFERENCES

- [88] E. Kyriacou, A. Kounoudes, L. Paraskeva, A. Constantinides, C. Pattichis, A. Jossif, M. Pattichis, and D. Vogiatzis, *Continuous monitoring of children with suspected cardiac arrhythmias*, Proc. of ehealth, 2009, pp. 1–8 (English).
- [89] E. Kyriacou, A. Nicolaides, C. Pattichis, S. Petroudi, M. Pattichis, M. Griffin, S. Kakkos, and G. Makris, *First and second order statistical texture features in carotid plaque image analysis: Preliminary results from ongoing research*, 33rd annual international conference of the IEEE Engineering in Medicine and Biology Society, IEEE EMBC, 2011Aug, pp. 6655–6658.
- [90] E. Kyriacou, C. Pattichis, M. Karaolis, C. Loizou, C. Christodoulou, M. Pattichis, S. Kakkos, and A. Nicolaides, *An integrated system for assessing stroke risk*, IEEE Engineering in Medicine and Biology Magazine **26** (2007Sept), no. 5, 43–50.
- [91] E. Kyriacou, C. Pattichis, and M. Pattichis, *An overview of recent health care support systems for emergency and mhealth applications*, 34th IEEE Annual International Conference of the Engineering in Medicine and Biology Society, 2009Sept, pp. 1246–1249.
- [92] E. Kyriacou, C. Pattichis, M. Pattichis, A. Jossif, L. Paraskeva, A. Constantinides, and D. Vogiatzis, *An m-health monitoring system for children with suspected arrhythmias*, Engineering in medicine and biology society, 2007. embs 2007. 29th annual international conference of the IEEE, 2007Aug, pp. 1794–1797.
- [93] E. Kyriacou, C. Pattichis, M. Pattichis, A. Jossif, D. Vogiatzis, L. Paraskeva, A. Constantinides, and A. Kounoudes, *An m-health system for continuous monitoring of children with suspected cardiac arrhythmias*, Proc. 4th European Conference of the International Federation for Medical and Biological Engineering, 2009, pp. 1325–1328.
- [94] E. Kyriacou, C. Pattichis, M. Pattichis, C. Loizou, C. Christodoulou, K. Kakkos, and A. Nicolaides, *A review of noninvasive ultrasound image processing methods in the analysis of carotid plaque morphology for the assessment of stroke risk*, IEEE Transactions on Information Technology in Biomedicine **14** (2010July), no. 4, 1027–1038.

## REFERENCES

- [95] E. Kyriacou, C. Pattichis, M. Pattichis, A. Mavrommatis, S. Panagiotou, C. Christodoulou, Kakkos S., and A. Nicolaides, *Classification of atherosclerotic carotid plaques using gray level morphological analysis on ultrasound images*, Proc. of the 3rd ifip conference on artificial intelligence applications and innovations, 2006, pp. 737–744 (English).
- [96] E. Kyriacou, M. Pattichis, K. Anastasis, C. Pattichis, L. Paraskeva, A. Jossif, D. Hoplaros, and D. Vogiatzis, *A wireless system for monitoring of children with suspected cardiac arrhythmias*, Proc. of the 2nd international conference on pervasive technologies related to assistive environments, 2009, pp. 52.
- [97] E. Kyriacou, M. Pattichis, C. Christodoulou, C. Pattichis, S. Kakkos, M. Griffin, and A. Nicolaides, *Ultrasound imaging in the analysis of carotid plaque morphology for the assessment of stroke*, Plaque imaging: Pixel to molecular, 2005, pp. 241–275.
- [98] E. Kyriacou, M. Pattichis, C. Christodoulou, C. Pattichis, Kakkos S., and A. Nicolaides, *Multiscale morphological analysis of the atherosclerotic carotid plaque*, 27th annual international conference of the ieee engineering in medicine and biology society, 2005Jan, pp. 1626–1629.
- [99] E. Kyriacou, M. Pattichis, C. Pattichis, A. Mavrommatis, C. Christodoulou, S. Kakkos, and A. Nicolaides, *Classification of atherosclerotic carotid plaques using morphological analysis on ultrasound images*, Applied Intelligence **30** (2009), no. 1, 3–23.
- [100] E. Kyriacou, M. Pattichis, C. Pattichis, A. Panayides, and A. Pitsillides, *m-health e-emergency systems: Current status and future directions*, IEEE Antennas and Propagation Magazine **49** (2007Feb), no. 1, 216–231.
- [101] E. Kyriacou, S. Petroudi, C. Pattichis, M. Pattichis, M. Griffin, S. Kakkos, and A. Nicolaides, *Prediction of high-risk asymptomatic carotid plaques based on ultrasonic image features*, IEEE Transactions on Information Technology in Biomedicine **16** (2012Sept), no. 5, 966–973.

## REFERENCES

- [102] E. Kyriacou, P. Vogazianos, C. Christodoulou, C. Loizou, A. Panayides, S. Petroudi, M. Pattichis, M. Pantziaris, A. Nicolaidis, and C. Pattichis, *Prediction of the time period of stroke based on ultrasound image analysis of initially asymptomatic carotid plaques*, Engineering in medicine and biology society (embc), 2015 37th annual international conference of the ieee, 2015Aug, pp. 334–337.
- [103] E. Kyriacou, S. Voskarides, C. Pattichis, R. Istepanian, M. Pattichis, and C. Schizas, *Wireless telemedicine systems: a brief overview*, Proc. 4th international workshop on enterprise networking and computing in healthcare industry (healthcom), 2002.
- [104] A. Lambrou, H. Papadopoulos, E. Kyriacou, C. Pattichis, M. Pattichis, A. Gammerman, and A. Nicolaidis, *Assessment of stroke risk based on morphological ultrasound image analysis with conformal prediction*, Artificial intelligence applications and innovations, 2010, pp. 146–153.
- [105] ———, *Evaluation of the risk of stroke with confidence predictions based on ultrasound carotid image analysis*, International Journal on Artificial Intelligence Tools **21** (2012), no. 04.
- [106] S. Lee, M. Pattichis, and A. Bovik, *Foveated image/video quality assessment in curvilinear coordinates*, Int. workshop on very low bitrate video coding, 1998, pp. 189–192.
- [107] ———, *Rate control for foveated mpeg/h.263 video*, Proc. ieee int. conf. on image processing, 1998Oct, pp. 365–369 vol.2.
- [108] ———, *Foveated video compression with optimal rate control*, IEEE Transactions on Image Processing **10** (2001Jul), no. 7, 977–992.
- [109] ———, *Foveated video quality assessment*, IEEE Transactions on Multimedia **4** (2002Mar), no. 1, 129–132.
- [110] Li-Jia Li, Richard Socher, and Fei-Fei Li, *Towards total scene understanding Classification, annotation and segmentation in an automatic framework.*, Cvpr, 2009, pp. 2036–2043.

## REFERENCES

- [111] D. Llamocca, C. Carranza, and M. Pattichis, *Separable fir filtering in fpga and gpu implementations: Energy, performance, and accuracy considerations*, 21st conference on field programmable logic and applications (fpl), 2011Sept, pp. 363–368.
- [112] ———, *Dynamic multiobjective optimization management of the energy-performance-accuracy space for separable 2-d complex filters*, 22nd ieee international conference on field programmable logic and applications (fpl), 2012Aug, pp. 579–582.
- [113] D. Llamocca, V. Murray, Y. Jiang, M. Pattichis, J. Lyke, and K. Avery, *A scalable open-source architecture for real-time monitoring of adaptive wiring panels*, Journal of Aerospace Information Systems **11** (2014), no. 6, 344–358.
- [114] D. Llamocca and M. Pattichis, *Real-time dynamically reconfigurable 2-d filterbanks*, Ieee southwest symposium on image analysis interpretation, 2010May, pp. 181–184.
- [115] ———, *A dynamically reconfigurable pixel processor system based on power/energy-performance-accuracy optimization*, IEEE Transactions on Circuits and Systems for Video Technology **23** (2013March), no. 3, 488–502.
- [116] ———, *Dynamic energy, performance, and accuracy optimization and management using automatically generated constraints for separable 2d fir filtering for digital video processing*, ACM Trans. Reconfigurable Technol. Syst. **7** (2014dec), no. 4, 31:1–31:30.
- [117] ———, *A self-reconfigurable platform for the implementation of 2d filterbanks with real and complex-valued inputs, outputs, and filter coefficients*, VLSI Design **2014** (2014), 5.
- [118] D. Llamocca, M. Pattichis, and C. Carranza, *A framework for self-reconfigurable dcts based on multiobjective optimization of the power-performance-accuracy space*, 7th international workshop on reconfigurable communication-centric systems-on-chip (recosoc), 2012July, pp. 1–6.
- [119] D. Llamocca, M. Pattichis, and A. Vera, *Partial reconfigurable fir filtering system using distributed arithmetic*, International Journal of Reconfigurable Computing **2010** (2010).



## REFERENCES

- [120] D. Llamocca, M. Pattichis, and G. Vera, *A dynamic computing platform for image and video processing applications*, 42nd asilomar conference on signals, systems and computers., 2009Nov, pp. 412–416.
- [121] ———, *A dynamically reconfigurable parallel pixel processing system*, 19th ieee international conference on field programmable logic, 2009Aug, pp. 462–466.
- [122] ———, *A dynamically reconfigurable platform for fixed-point fir filters*, International conference on reconfigurable computing and fpgas, 2009Dec, pp. 332–337.
- [123] D. Llamocca, M. Pattichis, G. Vera, and J. Lyke, *Dynamic partial reconfiguration through ethernet link*, Aiaa infotech@ aerospace 2010, 2010.
- [124] C. Loizou, M. Morega, E Kyriacou, S. Pasca, S Petroudi, P. Bamidis, M. Pattichis, and C. Pattichis, *Guest editorial to special issue on biomedical monitoring technologies*, International Journal of Monitoring and Surveillance Technologies Research (IJMSTR) (2013).
- [125] C. Loizou, V. Murray, M. Pattichis, M. Pantziaris, A. Nicolaidis, and C. Pattichis, *Am-fm texture image analysis and retrieval of the intima and media layers of the carotid artery*, 19th international conference on artificial neural networks, 2009.
- [126] ———, *Despeckle filtering for multiscale amplitude-modulation frequency-modulation (am-fm) texture analysis of ultrasound images of the intima-media complex*, International journal of biomedical imaging **2014** (2014).
- [127] C. Loizou, V. Murray, M. Pattichis, M. Pantziaris, and C. Pattichis, *Am-fm texture image analysis in brain white matter lesions in the progression of multiple sclerosis*, Proc. ieee southwest symposium on image analysis interpretation, 2010May, pp. 61–64.
- [128] ———, *Multiscale amplitude-modulation frequency-modulation (am-fm) analysis of ultrasound images of the intima and media layers of the carotid artery*, IEEE Transactions on Information Technology in Biomedicine **15** (2011March), no. 2, 178–188.

## REFERENCES

- [129] C. Loizou, V. Murray, M. Pattichis, M. Pantziaris, I. Seimenis, and C. Pattichis, *Am-fm texture image analysis in multiple sclerosis brain white matter lesions*, Proc. of ifmbe, 2010, pp. 446–449.
- [130] C. Loizou, V. Murray, M. Pattichis, I. Seimenis, M. Pantziaris, and C. Pattichis, *Multiscale amplitude-modulation frequency-modulation (am-fm) texture analysis of multiple sclerosis in brain mri images*, IEEE Transactions on Information Technology in Biomedicine **15** (2011Jan), no. 1, 119–129.
- [131] C. Loizou, M. Pantziaris, A. Nicolaides, A. Spanias, M. Pattichis, and C. Pattichis, *Ultrasound imaging media layer texture analysis of the carotid artery*, 8th ieee international conference on bioinformatics and bioengineering, 2008Oct, pp. 1–6.
- [132] C. Loizou, M. Pantziaris, M. Pattichis, E Kyriacou, and C. Pattichis, *Ultrasound image texture analysis of the intima and media layers of the common carotid artery and its correlation with age and gender*, Computerized Medical Imaging and Graphics **33** (2009), no. 4, 317–324.
- [133] C. LópezLeiva, S. Celedón-Pattichis, and M. Pattichis, *Participation in the advancing out-of-school learning in mathematics and engineering (aolme) project: Supporting middle school latinas bilingual and stem identities*, Girls and women of color in stem: Navigating the double bind., 2015.
- [134] C. LopezLeiva, S. Celedon-Pattichis, and M. Pattichis, *Participation in the advancing out-of-school learning in mathematics and engineering (aolme) project: Supporting middle school latinas' bilingual and stem identities*, Girls and women of color in stem: Navigating the double blind, 2016. Accepted for publication.
- [135] C. LopezLeiva, S. Celedon-Pattichis, M. Pattichis, and J. Morales Flores, *Teaching and learning binary numbers: Underrepresented students accessing high-quality stem practices.*, Access and equity: Promoting high-quality mathematics in grades 6-8. reston, in press.

## REFERENCES

- [136] C. LopezLevia, S. Celedon-Pattichis, and M. Pattichis, *Integrating mathematics, engineering and technology through mathematics modeling and video representations*, 13th international congress on mathematical education, 2016.
- [137] Zelun Luo, Boya Peng, De-An Huang, Alexandre Alahi, and Li Fei-Fei, *Unsupervised learning of long-term motion dynamics for videos*, 2017.
- [138] G. Martin and M. Pattichis, *The characterization of scanning noise and quantization on texture feature analysis*, Proc. 6th IEEE Southwest Symposium on Image Analysis and Interpretation, 2004 March, pp. 152–156.
- [139] U. Meyer-Baese, A. Vera, A. Meyer-Baese, M. Pattichis, and J. Perry, *An undergraduate course and laboratory in digital signal processing with field programmable gate arrays*, IEEE Transactions on Education **53** (2010 Nov), no. 4, 638–645.
- [140] U. Meyer-Baese, G. Vera, A. Meyer-Baese, M. Pattichis, and R. Perry, *Discrete wavelet transform fpga design using matlab/simulink*, Proc. SPIE Independent Component Analysis, Wavelets, Unsupervised Smart Sensors, and Neural Networks, 2006, pp. 624703–624703–10.
- [141] ———, *Smart Altera firmware for DSP with FPGAs*, Proc. SPIE Defense & Security Symposium, 2007, pp. 65760T–65760T–11.
- [142] ———, *DSP with FPGAs: a Xilinx/Simulink-based course and laboratory*, Proc. SPIE, 2008, pp. 697907–697907–12.
- [143] U. Meyer-Baese, G. Vera, S. Rao, K. Lenk, and M. Pattichis, *FPGA wavelet processor design using language for instruction-set architectures (LISA)*, Proc. SPIE Defense & Security Symposium, 2007, pp. 65760U–65760U–12.
- [144] S. Murillo, V. Murray, C. Loizou, C. Pattichis, M. Pattichis, and S. Barriga, *Motion and deformation analysis of ultrasound videos with applications to classification of carotid artery plaques*, SPIE Medical Imaging, 2012.
- [145] S. Murillo and M. Pattichis, *Motion estimation of carotid artery plaques*, Ultrasound and Carotid Bifurcation Atherosclerosis, 2012, pp. 355–378.

## REFERENCES

- [146] S. Murillo, M. Pattichis, and S. Barriga, *A review of motion estimation methods for non-invasive ultrasound motion and emerging strain imaging methods of carotid artery plaques*, International Journal of Experimental and Computational Biomechanics **1** (2011), no. 4, 359–380.
- [147] S. Murillo, M. Pattichis, C. Loizou, E. Kyriacou, A. Constantinides, and A. Nicolaides, *Atherosclerotic plaque motion analysis from ultrasound videos*, 40th asilomar conference on signals, systems and computers, 2006Oct, pp. 836–840.
- [148] S. Murillo, M. Pattichis, C. Loizou, C. Pattichis, E. Kyriacou, A. Constantinides, and A. Nicolaides, *Atherosclerotic plaque motion trajectory analysis from ultrasound videos*, Cd-rom proc. of the 5th international ieee embs special topic conference on information technology in biomedicine, 2006, pp. 1–5.
- [149] S. Murillo, M. Pattichis, P. Soliz, S. Barriga, C. Loizou, and C. Pattichis, *Global optimization for motion estimation with applications to ultrasound videos of carotid artery plaques*, Proc. spie medical imaging: Computer-aided diagnosis, 2010, pp. 76290X–76290X.
- [150] V. Murray, C. Agurto, S. Barriga, M. Pattichis, and P. Soliz, *Real-time diabetic retinopathy patient screening using multiscale am-fm methods*, 19th ieee international conference on image processing (icip), 2012Sept, pp. 525–528.
- [151] V. Murray, S. Barriga, P. Soliz, and M. Pattichis, *Survey of am-fm methods for applications on medical imaging*, Ibero-american conference on trends in engineering education and collaboration, 2009.
- [152] V. Murray, G. Feucht, J. Lyke, M. Pattichis, and J. Plusquellic, *Cell-based architecture for reconfigurable wiring manifolds*, Aiaa infotech@ aerospace conference, 2010.
- [153] V. Murray, D. Llamocca, J. Lyke, K. Avery, Y. Jiang, and M. Pattichis, *Cell-based architecture for adaptive wiring panels: A first prototype*, Cell **10** (2013), no. 4.

## REFERENCES

- [154] V. Murray, S. Murillo, M. Pattichis, C. Loizou, C. Pattichis, E. Kyriacou, and A. Nicolaides, *An am-fm model for motion estimation in atherosclerotic plaque videos*, 41st asilomar conference on signals, systems and computers, 2007Nov, pp. 746–750.
- [155] V. Murray and M. Pattichis, *Am-fm demodulation methods for reconstruction, analysis and motion estimation in video signals*, Ieee southwest symposium on image analysis and motion estimation in video signals, 2008March, pp. 17–20.
- [156] V. Murray, M. Pattichis, S. Barriga, and P. Soliz, *Recent multiscale am-fm methods in emerging applications in medical imaging*, EURASIP Journal on Advances in Signal Processing **2012** (2012), no. 1 (English).
- [157] V. Murray, M. Pattichis, H. Davis, S. Barriga, and P. Soliz, *Multiscale am-fm analysis of pneumoconiosis x-ray images*, 16th ieee international conference on image processing, 2009Nov, pp. 4201–4204.
- [158] V. Murray, M. Pattichis, D. Llamocca, and J. Lyke, *Field-programmable wiring systems*, Proceedings of the IEEE **103** (2015July), no. 7, 1159–1180.
- [159] V. Murray, M. Pattichis, and P. Soliz, *Analysis methods for retinal image characterization*, Proc. 42nd asilomar conference on signals, systems and computers, 2008, pp. 664–668.
- [160] ———, *Retrieval and classification of pneumoconiosis chest radiograph images using multiscale am-fm methods*, Proc. 43rd asilomar conference on signals, systems and computers, 2009, pp. 12–16.
- [161] ———, *Multiscale directional am-fm demodulation of images using a 2d optimized method*, 18th ieee international conference on image processing (icip), 2011Sept, pp. 249–252.
- [162] V. Murray, P. Rodriguez, M. Noriega, A. Dasso, and M. Pattichis, *2d amplitude-modulation frequency-modulation - based method for motion estimation*, Circuits systems (lascas), 2015 ieee 6th latin american symposium on, 2015Feb, pp. 1–4.

## REFERENCES

- [163] V. Murray, P. Rodriguez, and M. Pattichis, *Robust multiscale am-fm demodulation of digital images*, Ieee international conference on image processing, 2007Sept, pp. I-465-I-468.
- [164] ———, *Multi-scale am-fm demodulation and image reconstruction methods with improved accuracy*, IEEE Transactions on Image Processing **19** (2010May), no. 5, 1138–1152.
- [165] ———, *2d instantaneous frequency-based method for motion estimation using total variation*, Ieee globalsip, 2014Dec, pp. 1009–1013.
- [166] H. Nasrabadi, M. Pattichis, P. Fisher, A. Nicolaidis, M. Griffin, G. Makris, E. Kyriacou, and C. Pattichis, *Measurement of motion of carotid bifurcation plaques*, Ieee 12th international conference on bioinformatics bioengineering (bibe), 2012Nov, pp. 506–511.
- [167] M. Neofytou, C. Loizou, V. Tanos, M. Pattichis, and C. Pattichis, *Classification and data mining for hysteroscopy imaging in gynaecology*, 4th european conference of the international federation for medical and biological engineering, 2009, pp. 918–922.
- [168] M. Neofytou, C. Pattichis, M. Pattichis, V. Tonos, E. Kyriacou, S. Pavlopoulos, and D. Koutsouris, *Texture analysis of the endometrium during hysteroscopy: Preliminary results*, Proc. ieee 26th international conference of the ieee engineering in medicine and biology society, 2004Sept, pp. 1483–1486.
- [169] M. Neofytou, C. Pattichis, V. Tanos, M. Pattichis, and E. Kyriacou, *Quantitative analysis of hysteroscopy imaging in gynaecological cancer*, Handbook of research on advanced techniques in diagnostic imaging and biomedical applications, 2009, pp. 16.
- [170] M. Neofytou, C. Pattichis, V. Tonos, M. Pattichis, E. Kyriacou, and D. Koutsouris, *The effect of color correction of endoscopy images for quantitative analysis in endometrium*, 27th annual international conference of the ieee engineering in medicine and biology society, 2005Jan, pp. 3336–3339.

## REFERENCES

- [171] M. Neofytou, M. Pattichis, C. Pattichis, V. Tanos, E. Kyriacou, and D. Koutsouris, *Texture-based classification of hysteroscopy images of the endometrium*, 28th annual international conference of the IEEE Engineering in Medicine and Biology Society, 2006Aug, pp. 3005–3008.
- [172] M. Neofytou, V. Tanos, I. Constantinou, E. Kyriacou, M. Pattichis, and C. Pattichis, *Computer aided diagnosis in hysteroscopic imaging*, IEEE Journal of Biomedical and Health Informatics **PP** (2014), no. 99, 1–1.
- [173] M. Neofytou, V. Tanos, M. Pattichis, E. Kyriacou, C. Pattichis, and C. Schizas, *Color multiscale texture classification of hysteroscopy images of the endometrium*, Proc. of 30th annual international conference of the IEEE Engineering in Medicine and Biology Society, 2008Aug, pp. 1226–1229.
- [174] M. Neofytou, V. Tanos, M. Pattichis, C. Pattichis, E. Kyriacou, and S. Pavlopoulos, *Color based texture - classification of hysteroscopy images of the endometrium*, Proc. of the 29th annual international conference of the IEEE Engineering in Medicine and Biology Society, 2007Aug, pp. 864–867.
- [175] M. Neofytou, Vasilis Tanos, M. Pattichis, Efthymoulos Pattichis C. Kyriacou, and Dimitris Koutsouris, *A standardized protocol for texture feature analysis of endoscopic images in gynaecological cancer*, BioMedical Engineering OnLine **6** (2007), no. 1, 44.
- [176] C. Nguyen, J. Havlicek, G. Fan, J. Caulfield, and M. Pattichis, *Robust dual-band mwir/lwir infrared target tracking*, 48th asilomar conference on signals, systems and computers, 2014Nov, pp. 78–83.
- [177] F. Ourique, V. Licks, R. Jordan, and M. Pattichis, *Automatic segmentation of scanning electron microscopy images for molecular aggregation profiling*, Proc. IEEE 25th international conference of the IEEE Engineering in Medicine and Biology Society, 2003Sept, pp. 702–705 Vol.1.

## REFERENCES

- [178] A. Panayides, Z. Antoniou, V. Barberis, M. Pattichis, C. Pattichis, and E. Kyriacou, *Abdominal aortic aneurysm medical video transmission*, Ieee-embs international conference on biomedical and health informatics, 2012Jan, pp. 679–682.
- [179] A. Panayides, Z. Antoniou, Y. Mylonas, M. Pattichis, A. Pitsillides, and C. Pattichis, *High-resolution, low-delay, and error-resilient medical ultrasound video communication using h.264/avc over mobile wimax networks*, IEEE Journal of Biomedical and Health Informatics **17** (2013May), no. 3, 619–628.
- [180] A. Panayides, Z. Antoniou, M. Pattichis, and C. Pattichis, *The use of h.264/avc and the emerging high efficiency video coding (hevc) standard for developing wireless ultrasound video telemedicine systems*, Asilomar conference on signals, systems and computers (invited), 2012Nov, pp. 337–341.
- [181] A. Panayides, Z. Antoniou, M. Pattichis, C. Pattichis, and A. Constantinides, *High efficiency video coding for ultrasound video communication in m-health systems*, Ieee embc, 2012Aug, pp. 2170–2173.
- [182] A. Panayides, A. Constantinides, M. Pattichis, E. Kyriacou, and C. Pattichis, *Adaptive real-time hevc encoding of emergency scenery video*, Eai 4th international conference on wireless mobile communication and healthcare (mobihealth), 2014Nov, pp. 217–217.
- [183] A. Panayides, C. Loizou, M. Pattichis, E. Kyriacou, C. Shizas, A. Nicolaides, and C. Pattichis, *Ultrasound video despeckle filtering for high efficiency video coding in m-health systems*, Ciwsp workshop (in honor of the 70th birthday of prof. constantinides), 2013Jan, pp. 1–4.
- [184] A. Panayides, M. Pattichis, A. Constantinides, and C. Pattichis, *M-health medical video communication systems: An overview of design approaches and recent advances*, Ieee embc, 2013July, pp. 7253–7256.
- [185] A. Panayides, M. Pattichis, and E. Kyriacou, *Effective, real-time ultrasound video communications over hspa networks using despeckle filtering* (2014).



## REFERENCES

- [186] A. Panayides, M. Pattichis, C. Loizou, M. Pantziaris, A. Constantinides, and C. Pattichis, *An effective ultrasound video communication system using despeckle filtering and hevc*, IEEE Journal of Biomedical and Health Informatics **19** (2015March), no. 2, 668–676.
- [187] A. Panayides, M. Pattichis, and C. Pattichis, *Wireless medical ultrasound video transmission through noisy channels*, Proc. of 30th annual international conference of the IEEE Engineering in Medicine and Biology Society, 2008Aug, pp. 5326–5329.
- [188] ———, *Hevc encoding for reproducible medical ultrasound video diagnosis*, Asilomar conference on signals, systems and computers(invited), 2013Nov, pp. 1117–1121.
- [189] ———, *Mobile-health systems use diagnostically driven medical video technologies*, IEEE Signal Processing Magazine **30** (2013Nov), no. 6, 163–172.
- [190] A. Panayides, M. Pattichis, C. Pattichis, C. Loizou, and M. Pantziaris, *Wireless ultrasound video transmission for stroke risk assessment: Quality metrics and system design*, International workshop on video processing and quality metrics for consumer electronics (vpqm), 2013.
- [191] A. Panayides, M. Pattichis, C. Pattichis, C. Loizou, M. Pantziaris, and A. Pitsillides, *Robust and efficient ultrasound video coding in noisy channels using h.264*, 34th IEEE Annual International Conference of the Engineering in Medicine and Biology Society, 2009Sept, pp. 5143–5146.
- [192] ———, *Towards diagnostically robust medical ultrasound video streaming using h.264*, Biomedical Engineering, 2009, pp. 219–237.
- [193] ———, *Atherosclerotic plaque ultrasound video encoding, wireless transmission, and quality assessment using h.264*, IEEE Transactions on Information Technology in Biomedicine **15** (2011May), no. 3, 387–397.
- [194] A. Panayides, M. Pattichis, C. Pattichis, and A. Pitsillides, *A review of error resilience techniques in video streaming*, Proc. of ISYC, International Conference on Intelligent Systems and Computing: Theory and Application, 2006, pp. 39–48.

## REFERENCES

- [195] ———, *A tutorial for emerging wireless medical video transmission systems*, IEEE Antennas and Propagation Magazine **53** (2011April), no. 2, 202–213.
- [196] A. Panayides, M. Pattichis, C. Pattichis, C. Schizas, A. Spanias, and E. Kyriacou, *An overview of recent end-to-end wireless medical video telemedicine systems using 3g*, Proc. of 32nd annual international conference of the IEEE Engineering in Medicine and Biology Society, 2010Aug, pp. 1045–1048.
- [197] Andreas S Panayides, Constantinos S Pattichis, and Marios S Pattichis, *The promise of big data technologies and challenges for image and video analytics in healthcare*, 50th asilomar conference on signals, systems and computers, 2016, pp. 1278–1282.
- [198] C. Pattichis, P. Bamidis, C. Christodoulou, E. Kyriacou, D. Mitsis, M. Pattichis, and C. Pitris, *Guest editorial introduction to the special issue on biomedical signal processing and analysis selected papers from itab 2009*, Biomedical Signal Processing and Control **6** (2011), no. 3, 217–218.
- [199] C. Pattichis, C. Christodoulou, M. Pattichis, M. Pantziaris, and A. Nicolaides, *An integrated system for the assessment of ultrasonic imaging atherosclerotic carotid plaques*, Proc. IEEE Int. Conf. on Image Processing, 2001, pp. 325–328 vol.1.
- [200] C. Pattichis, E. Kyriacou, C. Christodoulou, M. Pattichis, C. Loizou, M. Pantziaris, and A. Nicolaides, *Cardiovascular: Ultrasonic imaging in vascular cases*, Wiley encyclopedia of biomedical engineering, 2006, pp. 12.
- [201] C. Pattichis, E. Kyriacou, M. Pattichis, A. Panayides, S. Mougiakakou, A. Pitsillides, and C. Schizas, *A brief overview of m-health e-emergency systems*, Proc. of the 6th international special topic conference on information technology applications in biomedicine, 2007Nov, pp. 53–57.
- [202] C. Pattichis, E. Kyriacou, M. Pattichis, A. Panayides, and A. Pitsillides, *A review of m-health e-emergency systems*, Cd-rom proc. of the 5th international IEEE EMBS special topic conference on information technology in biomedicine, 2006.

## REFERENCES

- [203] C. Pattichis, E. Kyriacou, S. Voskarides, M. Pattichis, R. Istepanian, and C. Schizas, *Wireless telemedicine systems: An overview*, IEEE Antennas and Propagation Magazine **44** (2002Apr), no. 2, 143–153.
- [204] ———, *Wireless telemedicine systems: An overview*, Telemedicine in practice, 2005, pp. 149–169.
- [205] C. Pattichis and M. Pattichis, *Time-scale analysis of motor unit action potentials*, Technical Report TR-97-11, Department of Computer Science, University of Cyprus, 1997.
- [206] ———, *Time-scale analysis of motor unit action potentials*, IEEE Transactions on Biomedical Engineering **46** (1999Nov), no. 11, 1320–1329.
- [207] ———, *Adaptive neural network imaging in medical systems*, Proc. of the 35th asilomar conference on signals, systems and computers, 2001, pp. 313–317.
- [208] C. Pattichis, M. Pattichis, and E. Micheli-Tzanakou, *Medical imaging fusion applications: An overview*, Proc. of the 35th asilomar conference on signals, systems, and computers, 2001Nov, pp. 1263–1267 vol.2.
- [209] C. Pattichis, M. Pattichis, and C. Schizas, *Wavelet analysis of motor unit action potentials*, Proc. of engineering in medicine and biology society, 1996Oct, pp. 1493–1495 vol.4.
- [210] C. Pattichis, C. Schizas, C. Kyriacou Efthymoulos, I. Fotiadis Dimitris, M. Pattichis, and D. Bamidis Panagiotis, *Guest editorial introduction to the special issue on citizen centered e-health systems in a global healthcare environment: Selected papers from itab 2009*, IEEE Transactions on Information Technology in Biomedicine **15** (2011), no. 1, 3–10.
- [211] C. Pattichis, C. Schizas, M. Pattichis, E. Micheli-Tzanakou, E. Kyriacou, and D. Fotiadis, *Guest editorial: Introduction to the special section on computational intelligence in medical systems*, IEEE Transactions on Information Technology in Biomedicine **13** (2009Sept), no. 5.

## REFERENCES

- [212] C. Pattichis, F. Schnorrenberg, C. Schizas, M. Pattichis, and K. Kyriacou, *A modular neural network system for the analysis of nuclei in histopathological sections*, Computational intelligence processing in medical diagnosis, 2002, pp. 291–322.
- [213] C. Pattichis, F. Schnorrenberg, N. Tsapatsoulis, C. Schizas, M. Pattichis, and E. Kyriacou, *A biopsy analysis support system for the detection and classification of breast cancer nuclei*, Emerging technology in breast imaging and mammography, 2008, pp. 441–462.
- [214] M. Pattichis, *Least squares fir filter design using frequency domain piecewise polynomial approximations*, 10th european signal processing conference, 2000Sept, pp. 1–4.
- [215] ———, *Novel algorithms for the accurate, efficient, and parallel computation of multidimensional, regional discrete fourier transforms*, 10th mediterranean electrotechnical conference(melecon), 2000, pp. 530–533 vol.2.
- [216] ———, *Future challenges and recommendations*, M-health: Emerging mobile health systems, 2006, pp. 365–370.
- [217] ———, *Section overview*, M-health: Emerging mobile health systems, 2006, pp. 273–276.
- [218] ———, *Multidimensional am-fm models and methods for biomedical image computing*, 34th ieee annual international conference of the engineering in medicine and biology society, 2009Sept, pp. 5641–5644.
- [219] M. Pattichis and F. Bostick, *Fast wavelet transform applications in electromagnetics*, Proc. of the int. conf. on digital signal processing, 1993.
- [220] M. Pattichis and A. Bovik, *Multi-dimensional frequency modulation in texture images*, International conference on digital signal processing, 1995, pp. 753–758.
- [221] ———, *A nonlinear fluid model for describing frequency modulation of image orientations*, Proc. ieee workshop on non-linear signal and image proc, 1995, pp. 198–201.

## REFERENCES

- [222] ———, *Am-fm expansions for images*, Proc. european signal processing conf., 1996, pp. 10–13.
- [223] ———, *A fluid model for texture images*, Proc. workshop on image and multidimensional signal processing, 1996.
- [224] ———, *Latent fingerprint analysis using an am-fm model*, Automatic fingerprint recognition systems, 2003, pp. 317–338.
- [225] ———, *Analyzing image structure by multidimensional frequency modulation*, IEEE Transactions on Pattern Analysis and Machine Intelligence **29** (2007May), no. 5, 753–766.
- [226] M. Pattichis, A. Bovik, J. Havlicek, and N. Sidiropoulos, *On the representation of wideband images using permutations for lossless coding*, Proc. of 4th iee southwest symposium image analysis and interpretation, 2000, pp. 237–241.
- [227] ———, *Multidimensional orthogonal fm transforms*, IEEE Transactions on Image Processing **10** (2001Mar), no. 3, 448–464.
- [228] M. Pattichis, T. Cacoullos, and P. Soliz, *Spatial image variability analysis*, 45th iee midwest symposium on circuits and systems, 2002Aug, pp. I–224–7 vol.1.
- [229] ———, *New models for region of interest reader classification analysis in chest radiographs*, Pattern Recognition, Special Issue of Digital Image Processing and Pattern Recognition Techniques for the Detection of Cancer **42** (2009), no. 6, 1058–1066.
- [230] M. Pattichis, S. Cai, C. Pattichis, and R. Abdallah, *An overview of digital video compression for mobile health systems*, M-health: Emerging mobile health systems, 2006, pp. 345–364.
- [231] M. Pattichis, C. Christodoulou, C. Pattichis, and A. Bovik, *Non-stationary texture classification using an am-fm model*, Proceeding sof international conference on neural networks, 1997Jun, pp. 1552–1557 vol.3.

## REFERENCES

- [232] M. Pattichis, J. Havlicek, S. Acton, and A. Bovik, *Multidimensional am-fm models with image processing applications*, Advances in image processing and understanding: A festschrift for thomas s huang, 2002, pp. 277–306.
- [233] M. Pattichis, V. Jatla, R. Hock, C. Henney, and C. Arge, *Detecting coronal holes for solar activity modeling*, 48th asilomar conference on signals, systems and computers, 2014Nov, pp. 89–93.
- [234] M. Pattichis, Y. Jiang, and D. Obregon, *System and methods for dynamic management of hardware resources*, Google Patents, 2015. US Patent 9,111,059.
- [235] M. Pattichis, H. Murallldharan, C. Pattichis, and P. Soliz, *New image processing models for opacity image analysis in chest radiographs*, 5th iee southwest symposium on image analysis and interpretation, 2002, pp. 260–264.
- [236] M. Pattichis and V. Murray, *Am-fm analysis over spatially bounded domains for applications in medical imaging*, 9th international conference on information technology and applications in biomedicine, 2009Nov, pp. 1–4.
- [237] M. Pattichis, G. Panayi, A. Bovik, and Shun-Pin Hsu, *Fingerprint classification using an am-fm model*, IEEE Transactions on Image Processing **10** (2001Jun), no. 6, 951–954.
- [238] M. Pattichis and C. Pattichis, *Fast wavelet transform in motor unit action potential analysis*, Proc. int. conf. of the iee engineering in medicine and biology society, 1993, pp. 1225–1226.
- [239] M. Pattichis, C. Pattichis, M. Avraam, A. Bovik, and K. Kyriacou, *Am-fm texture segmentation in electron microscopic muscle imaging*, IEEE Transactions on Medical Imaging **19** (2000Dec), no. 12, 1253–1258.
- [240] M. Pattichis, C. Pattichis, C. Christodoulou, D. James, L. Ketai, and P. Soliz, *A screening system for the assessment of opacity profusion in chest radiographs of miners with pneumoconiosis*, 5th iee southwest symposium on image analysis and interpretation, 2002, pp. 130–133.

## REFERENCES

- [241] M. Pattichis, H. Petropoulos, and W. Brooks, *Mri brain image segmentation using an am-fm model*, 34th asilomar conference on signals, systems and computers, 2000Oct, pp. 906–910 vol.2.
- [242] M. Pattichis, H. Pollard, J. Parra, A. Vera, C. Kief, and D. Bouldin, *Educational uses of fpgas, invited article describing unnm’s lab*, IEEE Circuits and Devices Magazine (2004Sept).
- [243] M. Pattichis, J. Ramachandran, M. Wilson, C. Pattichis, and P. Soliz, *Optimal scanning, display, and segmentation of the international labor organization (ilo) x-ray images set for pneumoconiosis*, Proceedings of the 14th ieee symposium on computer-based medical systems, 2001, pp. 511–515.
- [244] M. Pattichis, P. Soliz, C. Pattichis, D. James, and L Ketai, *Computer assisted morphological analysis of opacities on the international labor organization standard radiographs for the pneumoconioses*, American Journal of Respiratory and Critical Care Medicine (2002).
- [245] M. Pattichis and R. Zhou, *A novel directional approach for the scalable, accurate and efficient computation of two dimensional discrete fourier transforms*, Technical Report AHPCC2000-019, 2000.
- [246] M. Pattichis, R. Zhou, and B. Raman, *New algorithms for computing directional discrete fourier transforms*, Proc. ieee int. conf. on image processing, 2001, pp. 322–325 vol.3.
- [247] Marios S Pattichis, Sylvia Celedon-Pattichis, and Carlos LopezLeiva, *Teaching image and video processing using middle-school mathematics and the raspberry pi*, Ieee international conference on acoustics, speech and signal processing (icassp), 2017, pp. 6349–6353.
- [248] R. Paul and M. Pattichis, *Nested random phase sequence sets: A link between am-fm demodulation and increasing operators with application to cardiac image analysis*,

## REFERENCES

- Proc. ieee southwest symposium on image analysis and interpretation, 2004March, pp. 196–200.
- [249] Toby Perrett and Dima Damen, *Recurrent assistance: Cross-dataset training of lstms on kitchen tasks*, The ieee international conference on computer vision (iccv), 2017Oct.
- [250] S. Petroudi, I. Constantinou, M. Pattichis, C. Tziakouri, K. Marias, and C. Pattichis, *6th european conference of the international federation for medical and biological engineering: Mbec 2014, 7-11 september 2014, dubrovnik, croatia*, Springer International Publishing, Cham, 2015.
- [251] ———, *Evaluation of spatial dependence matrices on multiscale instantaneous amplitude for mammogram classification*, 6th european conference of the international federation for medical and biological engineering, 2015, pp. 156–159 (English).
- [252] S. Petroudi, I. Constantinou, C. Tziakouri, M. Pattichis, and C. Pattichis, *Investigation of am-fm methods for mammographic breast density classification*, 13th ieee international conference on bioinformatics and bioengineering, 2013Nov, pp. 1–4.
- [253] S. Petroudi, C. Loizou, M. Pantziaris, M. Pattichis, and C. Pattichis, *A fully automated method using active contours for the evaluation of the intima-media thickness in carotid us images*, 33rd annual international conference of the ieee engineering in medicine and biology society, ieee embc, 2011Aug, pp. 8053–8057.
- [254] C. Pizano, G. Heileman, C. Abdallah, and M. Pattichis, *Are perfect image watermarking schemes possible?*, 10th ieee mediterranean electrotechnical conference(melecon), 2000, pp. 669–672 vol.2.
- [255] J. Ramachandran, M. Pattichis, L. Scuderi, and J. Baba, *Tree image growth analysis using instantaneous phase modulation*, EURASIP Journal on Advances in Signal Processing **2011** (2011).



## REFERENCES

- [256] J. Ramachandran, M. Pattichis, and P. Soliz, *Pre-classification of chest radiographs for improved active shape model segmentation of ribs*, 5th iee southwest symposium on image analysis and interpretation, 2002, pp. 188–192.
- [257] J. Ramachandran, M. Pattichis, P. Soliz, and M. Wilson, *A hierarchical segmentation model for the lung and the inter-costal parenchymal regions of chest radiographs*, 45th iee midwest symposium on circuits and systems, 2002Aug, pp. I-439–42 vol.1.
- [258] K. Ramamurthy, A. Spanias, L. Hinnov, C. Akujuobi, M. Stiber, M. Pattichis, E. Doring, C. Pattichis, H. Thornburg, A. Papandreou-Suppappola, A. Spanias, R. Ayyanar, E. Campana, and S. Haag, *Work in progress - collaborative multidisciplinary j-dsp software project*, 39th asee/iee frontiers in education conference, 2009Oct, pp. 1–2.
- [259] N. Ray, J. Havlicek, S. Acton, and M. Pattichis, *Active contour segmentation guided by am-fm dominant component analysis*, Proc. iee int. conf. on image processing, 2001, pp. 78–81 vol.1.
- [260] Alexander Richard, Hilde Kuehne, and Juergen Gall, *Temporal action labeling using action sets*, CoRR [abs/1706.00699](https://arxiv.org/abs/1706.00699) (2017), available at [1706.00699](https://arxiv.org/abs/1706.00699).
- [261] ———, *Weakly supervised action learning with rnn based fine-to-coarse modeling*, 2017.
- [262] P. Rodriguez, V. Murray, and M. Pattichis, *A regularized optimization approach for am-fm reconstructions*, Proc. asilomar conference on signals, systems and computers, 2010Nov, pp. 219–221.
- [263] P. Rodriguez and M. Pattichis, *Adaptive sampling and processing of ultrasound images*, Proc. of the 35th asilomar conference on signals, systems and computers, 2001Nov, pp. 323–327 vol.1.
- [264] ———, *Real-time am-fm analysis of ultrasound video*, 45th iee midwest symposium on circuits and systems, 2002, pp. 216–219.

## REFERENCES

- [265] ———, *New algorithms for fast and accurate am-fm demodulation of digital images*, Ieee international conference on image processing, 2005Sept, pp. II-1294-7.
- [266] P. Rodriguez, M. Pattichis, and M. Goens, *M-mode echocardiography image and video segmentation based on am-fm demodulation techniques*, Proc. of the 25th international conference of the ieee engineering in medicine and biology society, 2003Sept, pp. 1176-1179 Vol.2.
- [267] P. Rodriguez, M. Pattichis, M. Goens, and R. Abdallah, *Object-based ultrasound video processing for wireless transmission in cardiology*, M-health: Emerging mobile health systems, 2006, pp. 491-508.
- [268] P. Rodriguez, M. Pattichis, and R. Jordan, *Computational simd framework: split-radix simd-fft algorithm, derivation, implementation and performance*, Proceedings of 14th international conference on dsp, 2002, pp. 861-864 vol.2.
- [269] ———, *Parallel single instruction multiple data (simd) fft: Algorithm and implementation*, Technical Report HPCERC2003-002, 2003.
- [270] Paul A Rodriguez-Valderrama, Marios Stephanou Pattichis, and Victor Manuel Murray-Herrera, *System and methods of regularized optimization for matrix factorization and image and video reconstruction*, Google Patents, 2014. US Patent 8,908,992.
- [271] F. Saqib, A. Dutta, J. Plusquellic, P. Ortiz, and M. Pattichis, *Pipelined decision tree classification accelerator implementation in fpga (dt-caif)*, IEEE Transactions on Computers **64** (2015Jan), no. 1, 280-285.
- [272] N. Sidiropoulos, M. Pattichis, A. Bovik, and J. Havlicek, *Coperm: Transform-domain energy compaction by optimal permutation*, Technical Report UT-CVIS-TR-97-002, Center for Vision and Image Sciences, The University of Texas at Austin, 1997.
- [273] ———, *Coperm: Transform-domain energy compaction by optimal permutation*, IEEE Transactions on Signal Processing **47** (1999Jun), no. 6, 1679-1688.

## REFERENCES

- [274] Karen Simonyan and Andrew Zisserman, *Two-stream convolutional networks for action recognition in videos*, Advances in neural information processing systems, 2014, pp. 568–576.
- [275] P. Soliz, H. Davis, V. Murray, M. Pattichis, S. Barriga, and S. Russell, *Toward automatic phenotyping of retinal images from genetically determined mono-and dizygotic twins using amplitude modulation-frequency modulation methods*, Spie medical imaging, 2010, pp. 76243F–76243F.
- [276] P. Soliz, C. Pattichis, M. Pattichis, D. James, and L. Ketai, *Texture analysis of opacity profusion in chest radiographs of miners with pneumoconiosis*, American Journal of Respiratory & Critical Care Medicine (2002).
- [277] P. Soliz, M. Pattichis, J. Ramachandran, and D. James, *Computer-assisted diagnosis of chest radiographs for pneumoconioses*, Proceedings of the spie medical imaging conference, 2001, pp. 667–675.
- [278] P. Soliz, S. Russell, M. Abramoff, S. Murillo, M. Pattichis, and H. Davis, *Independent component analysis for vision-inspired classification of retinal images with age-related macular degeneration*, Ieee southwest symposium on image analysis and interpretation, 2008March, pp. 65–68.
- [279] Nitish Srivastava, Elman Mansimov, and Ruslan Salakhudinov, *Unsupervised learning of video representations using lstms*, International conference on machine learning, 2015, pp. 843–852.
- [280] M. Thurgood, R. Jordan, and C. Abdallah, *Fast algorithms for generating the original and symmetric pons matrices*, 10th digital signal processing workshop, 2002.
- [281] A. Ulloa, J. Liu, V. Vergara, J. Chen, V. Calhoun, and M. Pattichis, *Three-way parallel independent component analysis for imaging genetics using multi-objective optimization*, 36th ieee embc, 2014Aug, pp. 6651–6654.

## REFERENCES

- [282] A. Ulloa, P. Rodriguez, J. Liu, V. Calhoun, and M. Pattichis, *A quasi-local method for instantaneous frequency estimation with application to structural magnetic resonance images*, 36th annual international conference of the ieee, embc, 2014Aug, pp. 1477–1480.
- [283] A. Vera, M. Pattichis, and J. Lyke, *A dynamic dual fixed-point arithmetic architecture for fpgas*, International Journal of Reconfigurable Computing **2011** (2011).
- [284] G. Vera, D. Llamocca, M. Pattichis, W. Kemp, W. Shedd, D. Alexander, and J. Lyke, *Dose rate upset investigations on the xilinx virtex iv field programmable gate arrays*, Ieee radiation effects data workshop, 2007July, pp. 172–176.
- [285] G. Vera, D. Llamocca, M. Pattichis, and J. Lyke, *A dynamically reconfigurable computing model for video processing applications*, Proc. 43rd asilomar conference on signals, systems and computers, 2009Nov, pp. 327–331.
- [286] G. Vera, U. Meyer-Baese, and M. Pattichis, *An fpga-based rapid prototyping platform for wavelet coprocessors*, Proc. spie defense & security symposium, 2007, pp. 657615–657615–10.
- [287] G. Vera, J. Parra, C. Kief, M. Pattichis, and L. Pollard, *Integrating reconfigurable logic in the first digital logic course*, Cd-rom proc. of icee conference, 2006.
- [288] P. Vieyres, G. Poisson, G. Triantafyllidis, M. Pattichis, and G. Kontaxakis, *Future challenges and recommendations on echography systems and services*, M-health: Emerging mobile health systems, 2006, pp. 509–514.
- [289] S. Voskarides, C. Pattichis, R. Istepanian, E. Kyriacou, M. Pattichis, and C. Schizas, *Mobile health systems: A brief overview*, Proceedings of spie aerosense, 2002, pp. 124–131.
- [290] Heng Wang and Cordelia Schmid, *Action recognition with improved trajectories*, Proceedings of the 2013 ieee international conference on computer vision, 2013, pp. 3551–3558.

## REFERENCES

- [291] Xiaofang Wang and Chun Qi, *Action recognition using edge trajectories and motion acceleration descriptor*, Machine Vision and Applications **27** (2016Aug), no. 6, 861–875.
- [292] H. Yu, S. Barriga, C. Agurto, S. Echegaray, M. Pattichis, W. Bauman, and P. Soliz, *Fast localization and segmentation of optic disc in retinal images using directional matched filtering and level sets*, IEEE Transactions on Information Technology in Biomedicine **16** (2012July), no. 4, 644–657.
- [293] H. Yu, S. Barriga, C. Agurto, S. Echegaray, M. Pattichis, G. Zamora, W. Bauman, and P. Soliz, *Fast localization of optic disc and fovea in retinal images for eye disease screening*, Spie medical imaging, 2011, pp. 796317–796317.
- [294] H. Yu, M. Pattichis, and M. Goens, *A robust multi-view freehand three-dimensional ultrasound imaging system using volumetric registration*, Proc. ieee international conference on systems, man and cybernetics, 2005Oct, pp. 3106–3111 Vol. 4.
- [295] ———, *Multi-view 3d reconstruction with volumetric registration in a freehand ultrasound imaging system*, Proc. spie medical imaging, 2006, pp. 614706–614706–12.
- [296] ———, *Robust segmentation and volumetric registration in a multi-view 3d freehand ultrasound reconstruction system*, 40th asilomar conference on signals, systems, and computers, 2006Oct, pp. 1978–1982.
- [297] ———, *Robust segmentation of freehand ultrasound image slices using gradient vector flow fast geometric active contours*, Ieee southwest symposium on image analysis and interpretation, 2006, pp. 115–119.
- [298] Joe Yue-Hei Ng, Matthew Hausknecht, Sudheendra Vijayanarasimhan, Oriol Vinyals, Rajat Monga, and George Toderici, *Beyond short snippets: Deep networks for video classification*, Proceedings of the ieee conference on computer vision and pattern recognition, 2015, pp. 4694–4702.

## REFERENCES

- [299] Bowen Zhang, Limin Wang, Zhe Wang, Yu Qiao, and Hanli Wang, *Real-time action recognition with enhanced motion vector cnns*, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2718–2726.
- [300] Luowei Zhou, Chenliang Xu, and Jason J Corso, *Procnets: Learning to segment procedures in untrimmed and unconstrained videos*, arXiv preprint arXiv:1703.09788 (2017).