Research & Creativity Portfolio

Dylan W. Pallickara

I stuttered till I was nine. During speech therapy, I was told that I was trying to get out too many things and that my words were tripping over each other. I needed to say what I wanted to in fewer words. Writing became an outlet for me. Being frugal with words helped fix my stutter and even affected my writing. I am drawn to frugal forms of expression: poetry, ASL, and computer programs. Whether it is poems or programs, I feel they can untangle mysteries and get us to a better place.

> Poems help me make sense of the world around me. Connections, changes, and the undercurrents of time that fuel them interest me deeply. Of course, a reader's own experience intertwines with and heightens some parts of the poem. In that sense, unlike prose, every reader's experience with a poem is like a fingerprint — unique.
>
> There is an elegance to program design and computational thinking that is very reminiscent of poetry. Exploring how a complex problem can be iteratively (or recursively) reduced into simpler problems – almost like fractals – before collating their intermediate results into a solution for arbitrary problem sizes is deeply satisfying.
>
> I find that there are similarities across poetry and programming. Fueled by rewrites and continuous edits, each attempt is limited only by my expressive power. Like the proverbial flapping of the butterfly's wings, small changes can have profound implications. A poem can meander or come unstuck. Similarly, differences in how I choose to express my program can substantially ramp up algorithmic complexity, making the problem intractable.

Finally, I find that each attempt – be it a poem or a program – is a time capsule encapsulating who I was and what I knew. Because I am an aspiring computer scientist and poet, I have put together a portfolio that reflects these interconnected aspects of whom I am. The first part is an abstract of my computing work and the second part is a few of my best poems.

How did I get to the point of embarking on a computer science research project? I started the journey of honing my computational thinking and programming skills while taking a few university and high-school courses: chronologically, (1) In the summer of 2020, I took an MIT EdX course on *Introduction to Computer Science and Programming Using Python*. (2) As part of my *AP Computer Science Principles* (9th grade, Academic Year 2020-21) course, my project was to provide an aggregation based on demographics and/or time of COVID cases – stored in a file – made publicly available by my county, and (3) In summer 2021, I completed the sophomore course on *Data Structures and Algorithms* offered by the Computer Science department at Colorado State University.

Research in Computer-Assisted Recognition of the American Sign Language

One million of the US population is deaf, and another 10 million have hearing disabilities [Mitchell, 2006]. Nine out of the ten children who are born deaf are born to parents who hear [NIDCD, 2023]. The inability to communicate with others feeds isolation and withdrawal from social situations. Non-verbal communications, using the American Sign Language (ASL), is the primary mechanism used by those who are deaf –and increasingly, those with hearing disabilities – to communicate. For those who are deaf, signing has benefits such as improved cognitive development, social engagement, and improved access to information [Foggetti, 2023].

The ability to communicate and decipher ASL is key to engaging socially and forming relationships with peers. However, if those who are verbal are unable to comprehend ASL it can be limiting. Because the average American interacts with 12 others over the course of a day [Zhaoyang, 2018], most interactions that individuals who are deaf are likely to have are with those who are verbal. The communication problem stems not from those who sign in ASL, but rather from those that cannot decipher ASL – roughly, 333 million individuals nationwide.

Over the past two years, I have worked on computer-assisted recognition of ASL signs and, more recently, on extensions to support instruction in ASL acquisition by identifying subtle errors in signing. Given the visual nature of ASL, the methods I explored were reliant on images. An alternative approach that has been explored by others involves the use of sensors [Zhou, 2020]. These require the ASL signer to wear gloves equipped with sensors that can be an onerous, expensive burden. Relying on images also allows non-signers to use their cellphone cameras – since over 97% of the US population has one [Pew Research, 2023], the barriers are reduced for everyone. I was also able to find an ASL dataset with a large number of images [Akash, 2017]. The availability of this corpus is what gravitated me to leveraging AI/ML (Artificial Intelligence/Machine Learning) methods that have shown demonstrable promise in learning from the data. The class of AI/ML methods that I have experimented with include Deep Neural Networks (DNNs) and Random Forests.

Challenges I had to circumvent few challenges to "train" models. All training and data processing had to be done on my personal laptop: a MacBook Pro. DNNs are computationally (and data) hungry; the bigger they are, the hungrier they get. Training models from scratch was infeasible – calibrating a DNNs large number of parameters entailed demands from a data and computational perspective that were impossible to satisfy. That set me off on a journey. I learned about *transfer learning* that allows a model trained on a different task to be repurposed for a new, related task. I leveraged open-source models that had previously been trained and calibrated for other tasks.

Being limited to using my personal laptop for all experimentation informed several of my choices. Rather than increase model complexity (and the accompanying data requirements), I decided to explore a more data-centric approach. I focused not so much on data volumes but rather on the qualitative aspects of the data such as reduced dimensionality and improved signal-to-noise in the individual datums used for training. I have found that smaller, well-curated datasets not only simplify the model structure, but that such datasets also reduce bias and improve accessibility all while increasing model accuracy. Finally, data curation allows an untethering of models from complex DNNs. The models can be based on traditional machine learning that are often substantially lighter weight with fewer parameters that need to be tuned.

Broad Description of the Work

My work in ASL can be broadly categorized into three chronologically progressive sections. Each built on lessons from the earlier sections. Failures and deficiencies in each section served as the north star for the next project.

1. ASL Raw Images: A lot of the effort here involved data wrangling, gaining experience with installing the TensorFlow deep learning framework, identifying a base model, and having a functional system. The accuracy of the ASL detection was poor.
2. ASL Wireframe: The next approach targeted distilling the ASL dataset further. This was accomplished via transformation of images with ASL hand signs into images with wireframes that encapsulate the joint structure of the metacarpals (palm) and phalanges (distal, intermediate, and proximate) that comprise the fingers. Crucially, this representation whittles away aspects such as skin tone, texture, or finger thickness and is effective in reducing bias: the model would be just as performant with skin tones and textures that were excluded from the original dataset. More importantly, this model was 94% accurate in the detection.

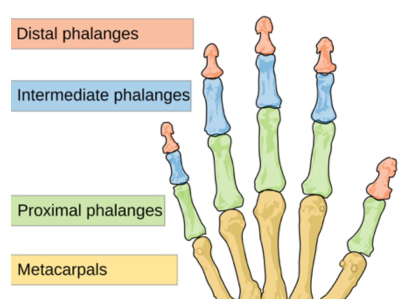

Figure 1: Bone in human hands. Source: Wikipedia

3. ASL Angles: Was there a way to further distill the information from the wireframe image? The wireframe image was transformed into a set of joint angles. The use of angles allows the detection to reconcile variations in the size of the palm (metacarpal) or the length of the fingers (phalanges). A consequence of this input distillation was that the classification model (now based on Random Forests) was less complex and much more interpretable. The model was also 97% accurate.

4. ASL Instructional Assistance: With a trove of ASL signs and their distilled representations, this project targeted ASL acquisition. To be useful, any instructional setting, must identify not just what's wrong, but also why and how it is wrong in a timely fashion. Targeted feedback identifies areas of improvement and when combined with deliberate practice can allow mastery of signs. When a user is trying to sign an ASL; the software first tries to detect that sign. Next, it identifies what were the key angular differences in the attempt. The model currently overlays the users attempt at a sign with the canonical representation of that sign (based on the averaging of joint angles in the reference dataset) and identifies angular differences.

ASL Detection (based on wireframes and angles) and ASL Instructional Assistance were performed under the mentorship and extended internship with Prof. Sarath Sreedharan in the Computer Science Department at CSU. I am presently the only student (graduate or undergraduate) who is doing work on ASL with Prof. Sreedharan.

What did the poet in me discover? My work with ASL was a petri dish for how networks "learn". Each input subtly adjusts the neural network's matrix of synaptic weights. Each input alters the weights (or mutates the DNA) of the deep network and leaves a trace of itself for posterity. I also saw how large language models (or LLMs) are exploitative of creative work by artists by "scraping content". Companies such as Google, OpenAI etc. display transparency for the structure of their networks; the data used to train their networks is a closely guarded secret.

I feel that the discourse around AI —with its emphasis on singularity and sentience — is disconnected from reality. The harms be it perpetuating bias, copyright violations, exploitation of creative work, are all here (now!) with little recourse to redress them. This exploitation of creative work will be perpetuated because newer versions of AI models, use the weights of earlier versions as the starting point. I want to research methods that force AI models to divulge works they have "scraped" and the bias/inequities they are inadvertently perpetuating.

ASL Scope: In my projects, the ASL sign detection is restricted to the numbers 0 through 9 as well as every alphabet except for "J" and "Z" that are not amenable for classifications using still images. The code and datasets have all been made publicly available via GitHub and Kaggle respectively (links on page-5). I am unfamiliar with the mobile app development process, so the models have not been made available in a mobile app — a pre-requisite for traction.

ASL Recognition with Raw Images

As part of my 10th grade International Baccalaureate (IB) personal project, I attempted to perform real-time detection of ASL gestures and translate the signs into text information in real time. A key goal of the IB personal project is to assess students' approaches to learning skills for self-management, research, communication, critical and creative thinking, and collaboration.

The objective of my IB project was to teach a computer to see and comprehend ASL to bridge the communication gap between ASL and verbal communicators. Without software-based assists, bridging the communication barriers would be possible only with translators. A drawback of human-mediated translations is not just the cost but also scale – there should be at least as many translators as the number of signers. A computer program and an affordable camera is cost-effective and significantly reduces the number of human translators that are needed.

I have built an application that takes 24 different ASL (American Sign Language) gestures and translates them to text information. As required by the IB program, I worked solo on this project in my 10th grade.

I had found a dataset of ASL images on Kaggle [Akash, 2017]. Rather than sift through the images of ASL gestures pixel-by-pixel and then try to codify rules by hand and enforce them in code, I figured using ML to codify these rules would be far more effective. Google's open-source TensorFlow supports several model architectures. I used the MobileNetv2 model that is part of TensorFlow's model garden and one among several models that Google makes available. MobileNetv2 that is a type of convolutional neural network designed for mobile and embedded vision applications. This architecture uses depth-wise, separable convolutions to build more compact deep neural networks and I leverage it in my project given its suitability for mobile and embedded devices.

This was my first attempt at fitting models to high dimensional data. For a while it seemed I was operating in the realm of Murphy's Law. There was a lot of trial-and-error. Most of the times things simply did not work. But as I started to work on it more, I found that my errors were starting to be much more rewarding. The errors started to point to new paths for experimentations. I learned how to capture images using the Open CV library. I was also able to create a virtual environment to protect the rest of my computer from software changes. I also learned, firsthand, the benefits of transfer learning and weight matrix initializations; I could use a model trained for a different task for a completely different objective.

ASL Detection (based on wireframes and angles) and ASL Instructional Assistance were performed under the mentorship and extended internship with Prof. Sarath Sreedharan in the Computer Science Department at CSU. I am presently the only student who is doing work on ASL with Prof. Sreedharan.

ASL Wireframes

Several efforts have targeted ASL recognition from raw images directly. However, depending on the data used to train the models, such methods may be subject to biases stemming from variations in the skin tone, texture, and finger thickness that may not have been sufficiently captured in the datasets (John, Sherif 2022). ASL recognition involved two phases. In the first phase, wireframes of hands were extracted while the second phase classifies the ASL sign based on the wireframe.

The primary dataset comprising images of ASL hand gestures is from Kaggle [Akash, 2017]. These images were transformed to extract wireframes associated with each gesture. The data are first passed through TensorFlow's hand landmarking functionality, which results in a set of 20 coordinates per image. A wireframe image (rendered over a black background) is generated from these coordinates as depicted in Figure 2. The generated wireframe image attenuates background interference and noise. Distilling each image into a simple wireframe also reconciles differences in skin tone, finger thickness, and skin texture.

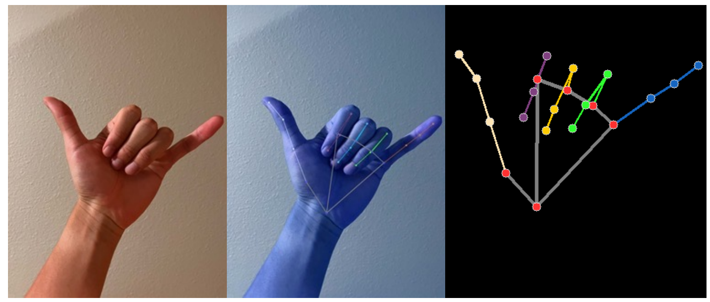

Figure 2: **(a)** Depiction of the ASL sign Y by me, and **(b)** the wireframe computed from (a) and superimposed on it. **(c)** the wireframe, rendered against a black background, computed from the letter Y for one of the images in the training dataset.

The wireframe extraction process was performed for 3,000 images per ASL sign to construct a curated dataset of wireframe images. This curated ASL dataset contained numbers 0 through 9 as well as every alphabet except for "J" and "Z" that are not amenable for classifications using still images. The curated ASL wireframe dataset was then used to train the deep network, which was a type of convolutional neural network. The 8-layer Keras Sequential Model was adapted and trained using TensorFlow. The model includes 4 convolutional input layers (2x32 feature maps with a size of 3x3 and 2x64 feature maps with a kernel size of 3x3). Two max pooling layers followed by a flatten layer were also used. Unlike the original implementation, a fully connected layer with 128 units was used. Dropout regularization was used and the overall dropout rate was set as 25% except for the last layer where the rate was set at 50%. The curated dataset was partitioned into training and validation datasets using an 80:20 split. Accuracy and loss metrics of the model were profiled at the end of training and the model had achieved a 94% accuracy (see Figure 3).

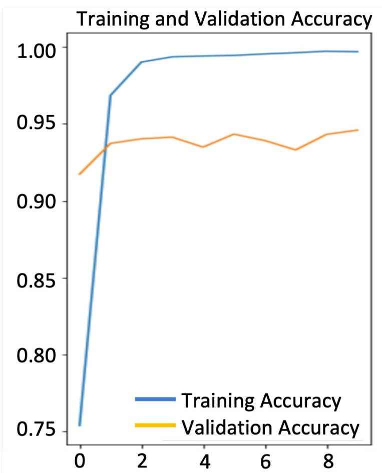

Figure 3: Classification accuracy for the ASL-WireFrames model.

Public release of the dataset and models for ASL Wireframes and ASL Joint Angles:

Both models and the Jupyter notebook have been publicly released on GitHub and are available for download and adaptation at https://github.com/Dylan-pallickara. The wireframe dataset for ASL has now been published on Kaggle (https://www.kaggle.com/datasets/dylanpallickara129/asl-alphabet-joint-angles) and is available for download and experimentation. Perhaps others can design more sophisticated models, expand on this data.

ASL-Joint Angles

Could the wireframe data be distilled even further? Each wireframe image was distilled into a set of 19 joint angles. The transformation process now involved two phases: conversion of raw images into wire frames, and wire frames into joint angles. Taking three data points (each normalized with a reference to the wrist or "point 0"), the angle of individual fingers relative to each other was calculated. The use of joint angles added an additional dimension, the classifications would now be scale-invariant i.e., independent of the length of the fingers.

Because the input space was now so compact, using a simpler model fitting algorithm (Random Forests) was now feasible. As the name suggests Random Forests includes a forest of decision trees. Incidentally, the decision thresholds within the individual trees are not random but are carefully chosen to maximize the information gained by the separation boundaries. One advantage of the Random Forests model is that it is simpler, and because it is based on decision trees, the decision thresholds are well-suited for scrutiny and interpretability.

The Random Forests model was configured to have 100 decision trees and had an accuracy of 97%.

ASL Instruction

Could a performant model (97% accuracy) be used as the basis for instruction? Systems with the requisite precision can allow ASL learners to receive immediate and quantifiable feedback regarding their signing accuracy — helping learners refine their skills without a physical teacher needing to be present. Currently, the vast majority of AI-based ASL research is focused on translation (Bantupalli 2018). But helping users refine their signing skills is an area that is lacking in targeted feedback.

The distilled image representations and the various joint angles that they encapsulate can be used as a training tool to support language acquisition. To this end, standards for each ASL hand sign were established. To accomplish this, joint angles were extracted from each wireframe image in the curated training datasets. An average set of finger angles was established for each sign.

This dataset of average joint angles for each ASL sign serves as the reference for identifying deviations. The teacher-model aspect of ASL acquisition involves two phases. In the first phase, wireframes are constructed from raw images of hand gestures representing ASL signs. Alongside the generation of wireframe, all the joint angles are also computed. Second, the classifier is used to identify the sign being attempted by the user. The predicted ASL sign is then cross-referenced with the established finger angles standard for that sign. This is done by identifying deviations for each individual finger and the overall similarity of the inputted hand sign. To find the similarity, each finger angle in the inputted image was subtracted from the average finger angle and divided by the average joint angle to find the extent of deviation. Then, the data regarding the accuracy of each finger joint and the overall average accuracy of the inputted hand sign are returned and displayed. Overlaying the deviation from the canonical wireframe reference, provides timely and targeted feedback that should help with picking up the language.

Bibliography

- Akash. ASL Alphabet. Image data set for alphabets in the American Sign Language. 2017. https://www.kaggle.com/datasets/grassknoted/asl-alphabet
- Bantupalli, Kshitij, and Ying Xie. "American sign language recognition using deep learning and computer vision." 2018 IEEE International Conference on Big Data (Big Data).
- Fernanda Foggetti. The Benefits of Sign Language for Children with Hearing Loss. Available at: https://www.handtalk.me/en/blog/the-benefits-of-sign-language-for-children-with-hearing-loss/#:~:text=Communicating%20in%20Sign%20Language%20has,for%20children%20and%20their%20development . Accessed 2023.
- John, J. , Sherif, B. 2022. Hand Landmark-Based Sign Language Recognition Using Deep Learning. In Machine Learning and Autonomous Systems: Proceedings of ICMLAS 2021 (pp. 147-157). Singapore: Springer Nature .
- Mitchell RE. How many deaf people are there in the United States? Estimates from the Survey of Income and Program Participation. J Deaf Stud Deaf Educ. 2006 Winter;11(1):112-9. doi: 10.1093/deafed/enj004. Epub 2005 Sep 21. PMID: 16177267.
- National Institute on Deafness and Other Communication Disorders (NIDCD). Quick Statistics About Hearing. Available at: https://www.nidcd.nih.gov/health/statistics/quick-statistics-hearing. Accessed 2023.
- Pangestu, Y.; Heryadi, Y.; Suparta, W.; Arifin, Y. 2022. The Deep Learning Approach For American Sign Language Detection. 2022 IEEE Creative Communication and Inno-vative Technology (ICCIT), Tangerang, Indonesia, 2022, pp. 1-5 doi: 10.1109/ICCIT55355.2022.10118626.
- Pew Research Center. Mobile Fact Sheet. The vast majority of Americans – 97% – now own a cellphone of some kind. Available at: https://www.pewresearch.org/internet/fact-sheet/mobile/#:~:text=The%20vast%20majority%20of%20Americans,a%20cellphone%20of%20some%20kind. Accessed 2023.
- Ravikiran, J.; Kavi Mahesh.; Suhas Mahishi, R.; Dheeraj, S. Sudheender.; and Nitin V. Pujari. 2009. Finger detec-tion for sign language recognition. Paper presented at the International MultiConference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, Hong Kong, March 18 - 20.
- Zhaoyang R, Sliwinski MJ, Martire LM, Smyth JM. Age differences in adults' daily social interactions: An ecological momentary assessment study. Psychol Aging. 2018 Jun;33(4):607-618. doi: 10.1037/pag0000242. Epub 2018 Apr 30. PMID: 29708385; PMCID: PMC6113687.
- Zhou, Z., Chen, K., Li, X. et al. Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays. Nat Electron 3, 571–578 (2020). https://doi.org/10.1038/s41928-020-0428-6

A Sample of Poems

Dylan W. Pallickara, 17, is a senior at Poudre High School in Fort Collins, Colorado. His work in poetry has appeared in The American Anthology of High School Poetry and The Dungeness Press. Dylan is an alumnus of the Iowa Young Writers' Studio and the Kenyon Review Young Writers Workshop. His writing has been recognized as a topical winner for the American Prize for High School Poetry and has been named a second-round finalist for the International Bridport Poetry Prize in the United Kingdom.

I am honored to share a small sample of my poems –a collection that represents my exploration of both the natural world and the human experience. I hope this collection can serve as a keyhole into my feelings and motivations.

"Puzzle Pieces" embodies the feeling of things being slightly out of reach.

In "Fingerprints," I grapple with the concept of self-discovery. The poem delves into the struggle of understanding one's identity, voice, and perception in a world that often seems elusive and disjoint from personal thought. It reflects on the challenge of truly knowing oneself and the impact of our interactions with the world.

"Jazz Standard" captures the essence of music and its ability to transcend time and space through the lens of an abecedarian poem i.e., each line starts with a consecutive alphabetical letter. The poem paints a picture of jazz tunes filling the air –intertwining with the sound of raindrops on rooftops and open doors. It explores the comfort found in the rhythm of life, the nostalgia of old songs, and the endless pursuit of coziness through musical melody.

In "Beauty Dies," I contemplate the fleeting nature of life and the inevitable passage of time. Silky flowers and changing seasons reflect the bittersweet essence of impermanence as the poem explores the cyclical nature of life –where hope gives way to realism, and dreams transform into memories.

Lastly, "Fall" captures the essence of nature's beauty amidst change. It tries to capture the transformation of seasons –rain-soaked trees and contorted clouds harbingers of discomfort. The poem reflects on the simultaneous melancholy and beauty found in nature's cyclical patterns, symbolizing the inevitability of change and the resilience of life.

I hope these poems resonate with you as much as they have resonated with me during their creation. Each piece represents a facet of the human experience, a fragment of emotion, and a glimpse into the depths of the soul. Thank you for taking the time to explore my work!

Warmly,

Dylan

Puzzle Pieces

I don’t know when I forgot how to solve puzzles
I remember the pieces fitting together so perfectly

the notches slipping
the colors blending

it’s so hard now
to hold myself together
I feel like my notches have split off
and my pieces don’t fit

each broken notch letting the wood of a table sneak through
bare pine where there should be color

*_Puzzle Pieces_ was a second-round finalist (<9% of submissions) for the 2023 Brid Port Poetry Prize
An International Poetry Competition in the UK

Fingerprints

I can't see myself
with my hands in my pockets
I've grown so used to the fabric on my ;
back, legs
they might as well be air

I can't see myself without
a reflection

Without : a freshly cleaned window
A newborn puddle, a rain
drop
the broken mirror of a misty road

How can you know someone you can't see?

I've grown used to my voice
-- mind tuning it out like the dull drone
of a midwestern wind

I only hear myself when I whisper
;strangely;
whispers carry further than words

I can only hear in a cave
maybe empty stairwells
convenience stores at 2:00 AM

and that summer silence that makes it feel like
you're wearing earmuffs
skin extending in tired breeze

I first heard my voice
through a musty speaker
a cluttered elementary classroom

I hate microphones

How can you know someone you can't see?
I can't recognize their voice ;

sometimes my soul runs away
a monarch flirting with the pecan tree
or a word snagged on the tip of
a wrinkled tongue

I can't look in a mirror
and see myself

can't sing to a clear night sky
and expect a conversation

I can't look myself in the eye
only look through
I'm stuck behind glass

I'm stuck in a mirror
trapped in old recordings on cheap USB drives

caught between ponds and puddles

the bars of a framed photograph

My touch is lost
gone in the reflection
of dusty piano keys
fingerprints on sliding glass doors

** Fingerprints was published in *The American Anthology of High School Poetry* and was the Topical Winner in the "*My World*" category.

Jazz Standard

Autumn Leaves dance like freshly poured coffee
            bending spinning,
                        cracking
                  don't worry, the trees will be fine most
                  everything blends orange to gray;

fine endings for delicate earth   Love Letters as
            great owls drift in Misty rain caught in a
                  hollow happiness

                  it's just like the old songs say ; the
                  jazz standards that is
            komfort is found in swirling rivers
letting the waning
                  moon sing on half frozen lakes and the

need to dip Little Suede Shoes in a puddle

open a door
            pin it to a wall and
                  quench the thirst for Giant Steps
                  reaching, stretching, wishing for tomorrow

            search for some meaning in
                  thin CDs and
                  use the wind as a backing track for old
                  vinyl ; scratched by a
                        weathered needle ; the willowy pinging of the
            xylophone braids
                        your fingers with the
                                    zoolike tones of an open door

Beauty Dies

once silk flowers
wilt to the texture of old parchment
as the glass of water begins to harden

a tree falls in the woods
and every day
a crystal ball is dropped

childhood dreams build on
;blind hope;
slowly die and blend to
pessimistic expectations

while freshly clipped grass gets
swept up by an angry gust of wind
as the verdant undertones
of and old forest begin
to be replaced by the gray of stone

Aspirational desires die when
the leg gets broken
when eyes fall out

but seasons change and
black gives way to blue
and satin rubs to a mirror gloss

Fall

the clouds mad and disfigured, slick with rainfall

like a contortionist slipping on black ice
weeping silently

rain hits the canopy of torn trees
mid april showers shouldn’t bring tears

ripping and cutting with
a warm tenderness
as dewdrops and puddles are drawn home

pulled towards the river like
nesting salmon

water runs as the albatross flies
the river gets fuller
the greens get greener

thin brown bends towards fresh water ;
like dancers inspired by bonsai trees
who’s leaves ;
grasp iridescent mountains
by the river a tree is bent to a snake