Introduction

Throughout human history, weather has played an integral role in shaping our lifestyles across different parts of the world, both as a boon and a detriment to our livelihood. Rain invigorates the land and brings about crops, yet it can cause floods and consequently, loss of life. Sunny days are usually the time for recreation and work, while hail and sandstorms are too disruptive and dangerous for day-to-day operations. Weather patterns like these greatly influence our schedules each day, and thus having an educated foresight into the weather in the near future can not only prevent needless injury, but also give insight to schools and workplaces on whether to continue operating or not. With the rise of turbulent climate patterns compounded with an increase of human activity in recent times, the significance of weather detection has become very apparent for environmental safety and climate change observation.

The Middle East, namely the Kingdom of Saudi Arabia, is home to a usually humid and hot climate. Rarely is it rainy or windy except for a month or two near the end of the year. It is this infrequent change in weather conditions that catches the public by surprise in events of sudden rain, strong winds, and sandstorms. These scenarios not only abruptly jumble school and employee schedules but also pose a threat to traffic safety due to lack of coordination and sophisticated sewage systems. Moreover, human-dependent forecasts may either take too long or may not be foreseen precisely due to subtle signs and sudden changes in weather patterns. In order to complement such shortcomings and bolster detection rates, we propose to implement deep learning techniques in the form of a transformer vision model to recognize weather efficiently and accurately.

Related Work

There has been a steady surge of interest on weather recognition through computer vision in recent years, with convoluted neural networks being among the most popular implementations. Elhoseiny et al. (2015) trained a CNN consisting of 5 convolution/pooling layers and 3 fully connected layers for binary classification of images of the Weather Database (‘cloudy’ and ‘sunny’). Using an 80-20 training/testing split, the model achieved an 82.2% average test accuracy on 2,000 images. The same dataset was the subject of another study conducted by An et al. (2018), whose combination of AlexNet, ResNet, and a multi-class SVM attained a 90% average test accuracy for binary classification. Their model performed even better on 4-class classification however, achieving an approximate accuracy of 98-99%. Al-Haija et al. (2020) achieved similar results using a ResNet18 model with pretrained ImageNet weights, observing a 98.22% test accuracy on the ‘multi-class weather recognition dataset’ which consisted of 4 classes: ‘sunrise’, ‘shine’, ‘rain’, and ‘cloudy’. Ibrahim et al. (2019) extracted features of weather images using the ‘WeatherNet’ model, a CNN parallelly comprised of four modified ResNet50 models. Data augmentation and sampling techniques were applied onto the Image2Weather and Multi-class Weather Image datasets to ensure balanced representations of both daytime and nighttime images. The strong feature extraction abilities of the model enabled it to achieve a precision of 92.4% on their test set when differentiating between ‘Sunny’, ‘Cloudy’, ‘Rainy’, and ‘Foggy’ classes. A deeper level of weather classification was conducted by Xiao et al. (2021), who developed the ‘MeteCNN’, a VGG16-inspired CNN model, to train on images of 11 classes: ‘Hail’, ‘Rainbow’, ‘Snow’, ‘Lighting’, ‘Dew’, ‘Sandstorm’, ‘Frost’, ‘Smog’, ‘Rime’, and ‘Glaze’. Omitting the VGG16’s fully connected layers in favor of a global average pooling layer among other modifications, the MeteCNN model attained the highest precision rates for ‘Hail’, ‘Rainbow’, ‘lightning’, and ‘dew’ from 97-100% while performing the weakest for ‘Snow’ and ‘Glaze’ at 85% precision. The model achieved an average accuracy of 92.68%, surpassing established models like ResNet34, MobileNet, and VGG19.