

Rotorcraft Flight Information Inference from Cockpit Videos using Deep Learning

Hikmat Khan
Rowan University
Glassboro, New Jersey,
United States

Ghulam Rasool
Rowan University
Glassboro, New Jersey,
United States

Nidhal Carla Bouaynaya
Rowan University
Glassboro, New Jersey,
United States

Charles C. Johnson
Federal Aviation Administration
Atlantic City, New Jersey, United States

PLEASE NOTE: DO NOT INCLUDE EMAIL ON 1ST PAGE

ABSTRACT

As the premier agency for promoting and insuring aviation safety, the Federal Aviation Administration (FAA) continues to promote and highlight the importance of participating in aviation Flight Data Monitoring (FDM) programs to improve flight safety and operational efficiency. Indeed, recorder safety is one of the agency's top 10 most wanted list of safety improvements in 2017-2018. The FAA, National Transportation Safety Board (NTSB), and the United States Helicopter Safety Team (USHST) are strong proponents of recorder use. These organizations and other industry partners are working together to implement a helicopter safety enhancement that promotes the use of flight data recorders as a mechanism to reduce the helicopter fatal accident rate. However, despite these best efforts to reduce the fatal accident rate with this lifesaving technology, barriers to implementation exist. These include initial costs of flight data recorders which can range from 9,000 – 50,000, on average. These costs can be significant for small operators and they combine to prohibit the widespread adoption of FDM by the rotorcraft community. Thus, rotorcraft, in general, typically have a lower participation rate in FDM programs than other forms of aviation (i.e. commercial fixed-wing or part 121 airline operations). On the other hand, even small helicopter operators often have access to or the financial means to purchase one or more off-the-shelf video cameras, which can be mounted inside the cockpit. These cameras offer an alternative to traditional flight data recorders as well as a means to augment them with supplementary data not always available depending on the type of Flight Data Recorder (FDR) installed in the helicopter. On board video data offers several possibilities for improving safety including flight replay, as well as the ability to extract information from the external scene such as readings of instrument panel gauges. As part of our research approach, we analyzed video data from cameras recording the instrument panel and compared these values against ground truth data from a representative flight data recorder. These values formed the training dataset for our video analytic framework. To analyze this information, we first captured the gauge of interest (i.e. airspeed indicator, tachometer, engine oil temperature/pressure) in each frame of every video. The gauge readings, extracted from all videos, were subsequently fed to train a deep Convolutional Neural Network (CNN) using the FDR measurements as ground truth. We fine-tuned Resnet50 CNN models for airspeed, engine oil temperature/pressure, and tachometer gauges. These models obtained 78%, 89%, 89%, and 88% validation accuracy on airspeed, engine oil temperature/pressure, and tachometer gauges, respectively. These results demonstrate the feasibility of an inexpensive cockpit camera solution that would facilitate participation in FDM programs even for legacy helicopters that may otherwise require significant installation work. To further demonstrate the feasibility, we used the trained models to retrieve airspeed and engine oil values from the complete flight profile. We observed that the our models predicted trajectories for gauges closely follow the actual sensory values recorded by FDR. Such solution results in an effective flight data analysis tool as well as improved safety and operational efficiency of rotorcraft.

INTRODUCTION

The National Transportation Safety Board (NTSB) continues to promote and highlight the importance of participating in aviation Flight Data Monitoring (FDM) programs to improve the efficiency and safety of rotorcraft operations. Operators within the helicopter community participate in several of these programs throughout the globe, although participation is not as widespread as that of other communities (i.e. fixed wing commercial airlines). An accurate and effective analysis of flight data can help in determining the potential risk of an accident and empower the necessary changes in standard operating procedures to mitigate the risk of occurrence. To participate in FDM, an operator needs to install and equip the helicopter(s) in their fleet with a Flight Data Recorder (FDR), which is also popularly known as the Black Box. A FDR records the flight activities and state information and allows for trending and analysis of flight data.

Compared to scheduled Part 121 or 135 commercial air service, rotorcraft and general aviation aircraft have historically featured a higher fatal accident rate. In the rotorcraft community, accident rates average more than 10 times higher than that of the fixed wing commercial aircraft, see helicopter accidents fact sheet in (Ref. 1). According to the International Helicopter Safety Team (IHST), the accident rate in the commercial and non-commercial rotorcrafts industry are 2.28 and 5.29 per 100,000 flights hours, respectively, in 2014. Accidents comparison is provided in (Ref. 1) and in response to the fatal accident rate, the United States Helicopter Safety Team (USHST) adopted an approach favored by the Commercial Aviation Safety Team (CAST) and the General Aviation Joint Steering Committee (GAJSC) that seeks to implement a data-driven approach to reduce the fatal accident rate by focusing on the root causes and developing safety enhancements that the community can implement. The FDM was one of the key safety enhancements approved by the USHST and its goal is to encourage and promote the adoption and use of flight data recorders on helicopters.

So, if FDRs offer all of this potential to enhance safety, why are they not implemented at a greater scale within the rotorcraft community? One answer may be that there remain significant barriers to implementation of an FDM program. These include not only the technical skills to operate an FDR, but also all of the costs to acquire and install an FDR as well as the costs to utilize it as part of an overall FDM program. In addition, these devices can require technical expertise and special reading devices.

Conversely, inexpensive cockpit cameras offer an alternative to traditional flight data recorders as well as a means to augment them with supplementary data not always

available on an FDR. These cameras can be mounted in a way that records the instrument panel and provides information that can augment or replicate information provided by a traditional FDR. Properly placed cockpit videos do not require any modification to the helicopter communication or display systems (Ref. 2). This is not always the case with traditional FDR's, which usually require a Supplemental Type Certificate (STC) or Field Approval (FA) to install and operate the device in accordance with the Rotorcraft Flight Manual (RFM). The videos of the instrument panel can be processed analytically to retrieve flight state information and facilitate post flight data analysis. Other advantages of the cockpit cameras are the ease of installation and data retrieval but also crash survive ability; as the captured videos, and the associated analytics, can be stored remotely. The International Helicopter Safety Foundation (IHSF) has the zero accident vision that would only be accomplished by promoting the helicopter safety culture across the global helicopter community. The IHSF has formed with prime responsibility to enable the sharing the aviation data for analysis that could help in reducing the overall accidents/fatalities rate. According to the report of International Helicopter Safety Foundation (IHSF), using the combine data of 49 countries across different regions has helped to reduce the helicopter accidents by 6 percent in 2017, as compared to the prior years and the helicopter fatal accidents were reduced by 17 percent year over year. Collectively comparing the reduction in accidents and fatal accidents from 2014 to 2017, the total accidents reduced by 32 percent and fatal accidents reduced by 42 percent (Ref. 8). The IHSF has conducted a worldwide survey and analyzed more than 1000 helicopter accidents. Such survey helped IHSF to offer seven different areas that could potentially help in prevent the helicopter accidents in future, that includes safety management system(SMS), training, health and usage monitoring system(HUMS), flight data monitoring system(FDM), wire strike prevention system, recommended maintenance and usage of the night vision (Refs. 9, 10).

The concept of system safety was first introduced in aviation in 1940 (Ref. 3). According to the NTSB safety recommendation A-00-31, FAA should take actions to address the safety issues concerned with the lack of cockpit imagery information (Ref. 4). The cockpit video recorder plays the role of a secondary black box while the use of cameras in rotorcraft provide a non-contact identification of the flaws in rotorcraft operations. The FAA has compiled a guide that contains information on various analysis that need to be performed to ensure system safety (Ref. 7). Different Aviation research has attempted to utilize flight video to infer the flight state information. Cooke *et al.* (Ref. 5) discussed an approach that uses thermal images to detect the presence or absence of cracks in rotor blades, rotor hub and swashplate assembly during the rotorcraft flight. Modern rotorcraft have a complex system design and it is necessary to perform early

Presented at the Vertical Flight Society 75th Annual Forum & Technology Display, Philadelphia, Pennsylvania, May 13–16, 2019. Copyright 2019 by AHS - The Vertical Flight Society. All rights reserved.

inspection and flight data analysis of this design to ensure system safety. Hewitt *et al.* (Ref. 6) discussed the advantages of early system safety approaches. Shin *et al.* (Ref. 11) proposed a spatial clustering technique to estimate a helicopter’s attitude, such as the bank angle using cockpit image data. Although their research questions were more related to ours, their approach was based on traditional image processing techniques, such as clustering. Moreover, their algorithm was not designed to address the challenges of cockpit videos, such as sunlight distortion, gauges obstruction by the pilots and different environmental lighting conditions. The proposed deep learning approach, on the other hand, learns from real flight videos, and is able to account for the changing conditions during the flight. In a similar line of research, Kuo *et al.* (Ref. 12) adopted an image processing technique to recover the RPM and main torque values for rotorcraft Bell-206.

Retrieving flight state information from the instrument panel videos is a daunting task subject to several challenges. First, frame by frame analysis of the cockpit videos is a tedious task. Second, the image quality deteriorates due to the rotorcraft engine vibration, occlusion and varying ambient lighting conditions. Third, flight state information inference from analog indicators or gauges is an extremely time consuming task for the post flight data analysis operators and post-accident investigators. To solve these challenges, we propose a video analytics framework based on Convolutional Neural Networks (CNNs) that automates retrieval of the flight state information from cockpit instrument panel videos. Our trained CNN models are able to infer the airspeed, engine oil temperature/pressure, and tachometer gauges with 78%, 89%, and 88% validation accuracies, respectively. Our proposed approach provides the flight states information in a time efficient manner not only for post flight review or investigation but could also be implemented in real-time. The main contribution of this research is to establish a video analytics framework, based on deep learning, that processes the gauges and indicators on the instrument panel and retrieves the digitized values of the flight parameters from the cockpit videos. The performance of the proposed algorithm has been demonstrated, evaluated, and tested for the indicated airspeed, engine oil temperature/pressure, and tachometer gauges of the cockpit instrument panel for a series of flight videos of a Sikorsky S76 helicopter.

The rest of the paper is organized as follows: Section *Convolutional Neural Networks (CNN)* explains the basic building blocks of Convolutional Neural Networks and their working. Section *Methodology* describes the overall methodology that includes dataset analysis and experimental setup. The Results Section discusses the trained model prediction of a complete flight test video. Finally, the Conclusion Section summarizes the main findings of the paper and discusses future work.

CONVOLUTIONAL NEURAL NETWORKS (CNNs)

CNN-based algorithms have obtained state-of-the-art performance in computer vision tasks, such as image classification, object detection, image recognition and semantic segmentation (Refs. 13–15). The basic reason behind their record-shattering performance is not only the ability of CNNs to take advantage of the spatial coherence in images but also to learn hierarchical features directly from the input domain data without the need of domain knowledge or hand engineered features. The early layers of the network learn general features, such as edges, and subsequent layers learn more concrete features of the problem domain. The CNN hierarchical features enable the algorithm to adapt to another domain that does not have a large amount of training data available. All recent winners of the ImageNet competition are variations of the CNN model. In most recent ImageNet competitions, CNN-based models surpassed human level performance on image classification tasks (Refs. 13–15).

Convolution consists in convolving a signal or an image with kernels to obtain feature maps. There are four key ideas behind CNNs that take advantage of the properties of natural signals: local connections, shared weights, pooling and the use of many layers (Ref. 17). A CNN architecture begins with a convolutional layer that probes an array input with a number of convolutional kernels (or filter weights) to extract spatial information. The result is then passed through a nonlinearity, such as ReLU, and the output is a feature map (Ref. 17). The number of the generated feature maps depends on the number of convolutional filters used. Multiple convolutional layers, followed by nonlinearities, enable the network to learn features in increasingly complex hierarchy. The feature maps are followed by a max pooling layer. The main purpose of using the max pooling layer is to reduce the input dimensionality, mitigate the risk of overfitting and reduce the computational cost. The final feature maps are fully connected to every neuron in a fully connected layer. Finally, a softmax function is used for classification. Figure 1 presents the block diagram of a typical CNN architecture.

Transfer Learning:

In deep neural networks, appropriate weight initialization can reduce the convergence time. In particular, initializing the network weights with pre-trained models can serve as a good starting point given the vast compute and data resources needed to develop these models. This approach is called transfer learning and has proved effective in achieving optimum accuracy (Refs. 17, 19). The idea is that early convolutional layers extract general, low-level features that are applicable across images, such as edges, patterns, gradients, and the later layers identify specific features within an image, such as ears or wheels. In our

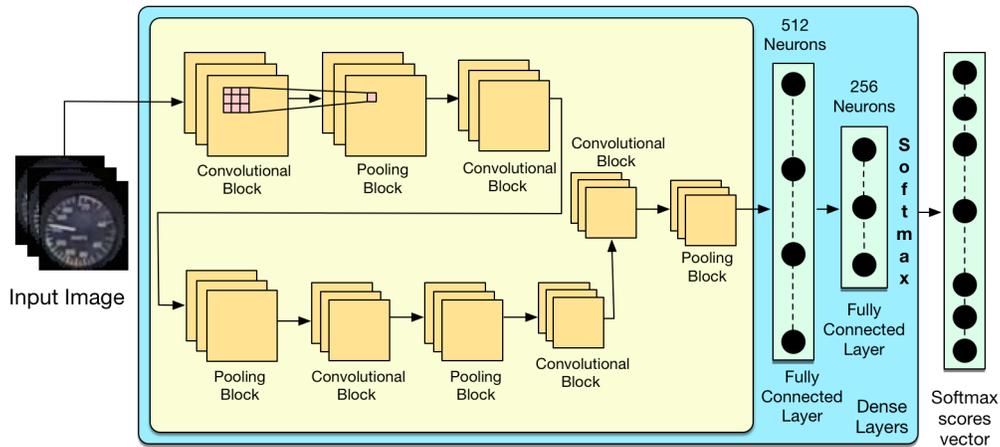


Fig. 1. Typical Convolutional Neural Network (CNN) architecture. Convolutional block convolve the image with kernel and output the feature maps. Pooling block is used to reduce the dimensionality of the feature maps. At fully connected layers, each pixel of the feature map is considered as a neuron and forwarded to the fully connected layers. The network architecture ends with a softmax vector that contains the softmax scores for individual classes.

experiments, we applied transfer learning and used a pre-trained model from ImageNet to initialize the weights.

Introducing nonlinearities:

The activation function introduces non-linearity into every layer of the network. There are many activation functions proposed in literature. Some of the most commonly used activation functions are: sigmoid, hyperbolic tangent, Rectified Linear Unit (ReLU), Leaky Rectified Linear Unit (LeakyReLU) and parametric ReLU (Ref. 17). Each activation function has some limitations and is not suitable for every scenario. ReLU obtained better results when compared with the sigmoid and hyperbolic tangent functions but the large gradient flowing through ReLU seems to update weights that will never activate at any data point (Ref. 17). Another issue with the ReLU activation function is that it ignores gradients smaller than zero. LeakyReLU is a different version of ReLU that tends to mitigate some of its issues by introducing negative gradients (Ref. 17).

Pooling Layers:

Pooling is used to reduce the dimensionality of the feature maps. It sub-samples feature maps by combining the spatially nearby features, and, subsequently, makes the model invariant to small intensity and illumination changes. The most commonly used pooling is max pooling, min pooling and average pooling (Ref. 17). The min and max pooling select features with the minimum and maximum value, respectively, in the pooling kernel. The average pooling calculates the average of the features in the pooling kernel.

Regularizations:

Regularization prevents the network from over fitting. There are several regularization schemes available: L_1 -Regularization, L_2 -Regularization, Batch Normalization, Global Average/Max Pooling, etc. L_1 -Regularization adds the magnitude of the weights/parameters as a penalty term to the loss function. L_2 -Regularization adds the squared magnitude of the parameters as a penalty term to the loss function. Batch normalization increases the stability of the network by normalizing the output of every activation layer by subtracting the batch mean and dividing by the batch standard deviation. Batch normalization allows each layer of the network to learn by itself a little bit more independently of other layers. Dropout is a simple yet effective technique for regularization that was recently proposed in (Ref. 17). During the training step, Dropout regularizer randomly switches on and off the neurons; thereby forcing each neuron to learn and contribute independently to the overall output of the network. Global average pooling performs a type of dimensionality reduction, thereby reducing overfitting.

Fully Connected Layers:

At the end of the convolution operations, the network has fully connected layers, where each pixel of the feature map is considered as a neuron and forwarded to every neuron in the fully connected layers. A classifier is used for the final classification step. The softmax function is the most commonly used classifier in deep neural networks.

Loss Function:

The loss function measures the compatibility between the predicted value of the network and the given ground truth label. The weights of a neural network are optimized, during the training phase, to minimize this loss function. For classification problems, the cross entropy loss is widely adopted; For regression problems, the L_2 or L_1 losses are used. For object detection problems, there is a loss function named *focal loss*, which is believed to ease the training process of the CNN based detector (Ref. 17). The most widely used loss functions: L_1 , L_2 and cross entropy are briefly described here. the L_1 loss is the sum of the magnitudes between the predicted labels and the true labels. It is mainly used for regression-type problems. The L_2 -loss is the sum of squared differences between the predicted labels and the true ones. The Cross Entropy loss, commonly used in classification, encodes the maximum likelihood estimation of the model (Ref. 17).

METHODOLOGY

Dataset

In this work, we used an FAA proprietary dataset of eleven instrument panel flight videos (i.e., Broomcloset view) recorded on Sikorsky S76 helicopters. The total duration of the eleven flight videos is around 15 hours. Figure 2 displays example gauges on the instrument panel. Table 1 presents the gauges training and validation set details. Figure 7 presents the class wise data distribution of training and validation examples for gauges.

Table 1. Overall Training and Validation Dataset

Gauge	Bin Size(α)	Training	Validation
Airspeed(kn)	1(kn)	227718	25363

Data Pre-processing: We performed a minimal level of pre-processing. The resolution of the gauges is very low, but we did not try to increase it. The CNN will learn appropriate features from this low resolution data and learn how to classify them. The only pre-processing we did is to extract the gauges from the videos. Specifically, we applied a circular mask that left only the desired gauge dial in the input image and zero multiplied the outside region, i.e., the remaining instrument panel and cockpit. Besides circular masking we did not perform any particular filtering, denoising or resolution enhancing of the given frames and gauges. The CNN learned the features associated with the given level of resolution and performed classification, see Fig.3 for an example of low resolution airspeed input images.

Extracting Gauges: CNN is a supervised machine learning algorithm. To generate training data for the CNN,

we processed flight instrument panel videos and annotated the individual frames with their corresponding FDR ground truth values. The FDR provided sensors' recording for each frame of the flight video. Each frame and corresponding FDR recordings have a unique time stamp that we leveraged to map each frame with its corresponding FDR readings. The FDR timestamp is embedded in the upper left corner position of each frame. Figure 4 presents the camera view that recorded the instrument panel and shows the embedded timestamp. However, there are few technical issues that needed to be solved beforehand. The FDR recorded the sensor readings with a fixed frequency of 10 readings per seconds (i.e., 10 Hz) while the corresponding flight video could have a varying frame capturing speed, i.e., frame per seconds (fps). Moreover, the FDR readings have missing sensor values at any particular time stamp. Because of the frequency matching and missing data issues, the dataset needed to be perfectly synchronized prior to training. To do this, we annotated the individual frames with FDR ground truth values following three steps. In the first step, we used the Optical Character Reading (OCR) algorithm (Ref. 16) to extract the time stamp of each frame. In the second step, we searched and selected the FDR sensor readings that matched the extracted frame timestamp. In the third step, we annotated the current frame with the selected FDR sensor readings. This way we managed to create the annotated dataset that could be used as a reliable training data for the supervised machine learning algorithm. The block diagram of this multi-step frame annotation procedure is shown in Fig. 5.

Thresholding Sensor Noise: We observed that FDR sensor readings below 35 kn for the airspeed gauge are noisy and unreliable. The noise is more prominent when the helicopter is in hovering state before landing or after take off. For instance, the analog needle of the airspeed gauge stays very close to 0 and does not correspond to the actual FDR sensor recorded airspeed. One possible cause is the sensitivity of FDR sensors and jittering maneuvers of the helicopter, especially when the rotorcraft is in hovering state. In order to generate reliable and clean training dataset for the supervised machine learning algorithm, we discarded all FDR values (and their corresponding frames) that indicated an airspeed lower than 35 kn. For the other gauges, we did not observe noise in FDR sensors readings, so we used the whole data.

Defining Bins for Classes: We used bin sizes to discretize the analog gauge values into digital bins. For instance, The maximum value in the airspeed gauge is 200 kn and the minimum is 0 kn, we divided the airspeed needle rotation into 200 bins (i.e. discrete steps) with a difference of 1 kn. An airspeed measurement between 40 kn and 41 kn is assigned to the class bin 41. For the airspeed classification problem, we used three bin sizes: 1 kn, 3 kn and 5 kn. For the the tachometer, N1/ NG1

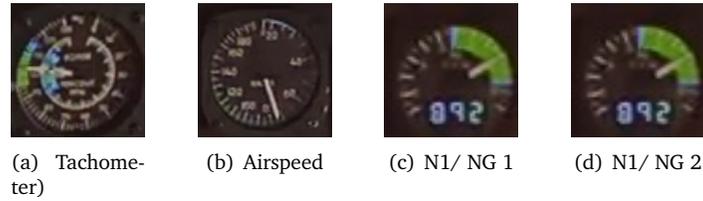


Fig. 2. Examples gauges on a helicopter instrument panel. From Left to Right: Tachometer, Airspeed, N1/NG1, N1/NG 2 gauges.



Fig. 3. 1st Row: Airspeed gauge images without glare, sun light distortion or varying ambient lights. 2nd Row: Airspeed gauges under challenging night light conditions. These frames have low readability even with an expert human eye. 3rd Row: Airspeed gauge images with different glaring and ambient effects.



Fig. 4. A sample camera view of the instrument panel that contains the different gauges. Each gauge displays the value of a flight parameter.

and N1/NG2 gauges, we used a bin size equal to 1. Although a bin size of 1 allows a finer classification of the sensor values, it presented some challenges for the airspeed gauge because of the noise discussed above. For instance, we observed that the needle position of the airspeed gauge does not move even though the FDR reading is indicating different values. Figure 6 shows the airspeed gauge with the needle in the same position (for the human eye), whereas the FDR readings are different. This means that, as we increase the bin size, the model will ob-

tain higher validation accuracy. In order to support this argument, we performed three experiments for airspeed with varying bin sizes of 1 kn, 3 kn and 5 kn. We observed that the model that trained with higher bin size has obtained higher validation accuracy as compared to the model trained on a smaller bin size dataset.

Gauges Data Distributions: For all gauges, the data across different bins are not uniformly distributed or balanced. Each gauge has its specific recommended operating interval throughout the flight. If the gauge is not within that recommended interval, then there is a fault either in the FDR sensor reading or an actual mechanical fault that needs to be addressed immediately to prevent hazardous situations. For example, N1/ NG1 and N1/ NG2 gauges measure the temperature of the helicopter engine, which is expected to remain in a specified interval throughout the flight. Other gauges, like the tachometer, jumps up to the interval 99-105, as soon as the rotary system has ignited. Similarly, for the airspeed, the pilots are required to maintain the ground speed of the rotorcraft between a maximum and minimum recommended speed limits. Thus, we can expect the airspeed needle to remain, most of the flight time, between some intervals on the gauge dial. This results in an imbalance distribution across the gauges. Table 1 presents the dataset information of different gauges, and Fig.7 presents the individual gauges bin-wise data distribution.

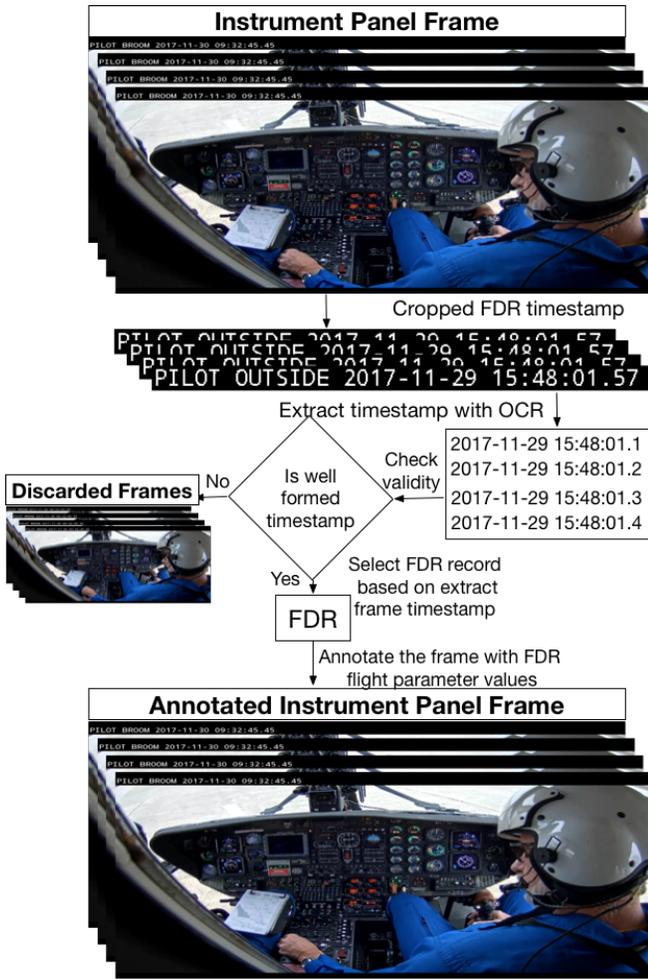


Fig. 5. A diagram representing the complete workflow to synchronize the individual frames of a flight video with their corresponding FDR sensor recordings based on time stamp matching.

Experimental Setup

Appropriate model selection is usually performed on a trial-error basis. A complex machine learning model may overfit the training data and perform poorly on the validation set, while a simple model may suffer from under fitting and will perform poorly on both the training and the validation sets. Hence, there is a trade off between model complexity and accuracy. Regularization has been shown to help balance complexity and accuracy. In our quest for an appropriate model, we experimented with different state-of-the-art CNN architectures, namely Resnet50 (Ref. 21), Inception (Ref. 23), Xception (Ref. 22) and VGG variants (Ref. 18). We found that Resnet50 achieved an optimum balance between (training and validation) accuracy and computational complexity. Resnet50 is a 50-layer CNN with residual connections that mitigate the risk of gradient vanishing during training. Although the architecture of Resnet50 has a complexity level comparable to Inception and Xception, it has an efficient implementa-

tion with a number of trainable parameters that is comparable to VGG variants. The more complex Inception and Xception models did not obtain significant improvements, in validation accuracy, over Resnet50. We used the following hyperparameters for all our experiments: the batch size is 96, the optimizer is *adam optimizer* with default parameter values as discussed in (Ref. 20), and a learning rate of 0.0001. The training was performed over 250 epochs. The same early stopping criteria of 7 epochs for maximum validation was used throughout all experiments. It is important to mention that we also experimented with transfer learning, i.e., freezing the earlier layers of the network and training the remaining layers only; but this strategy did not achieve a good accuracy. Therefore, we initialized the network with Imagenet weights and kept all the layers learnable.

RESULTS

Table 2 summarizes the validation accuracy of Resnet50 for airspeed, tachometer, N1/NG1 and N1/NG2 gauges. To demonstrate the prediction performance of the model, we tested it on two complete flight videos that the network never saw before. Figures 8(a) and 9(a) show the predicted airspeed trajectory vs the FDR recorded airspeed values. The first flight (Fig. 8) was flown in the day time while the second flight (Fig. 9) was flown in the night time. Both flight settings (i.e., day time and night time) have their own challenges, e.g., sun light distortion, different illumination condition, glaring, etc. Figure 3 shows examples of gauge lighting conditions, which makes the tracking of the analog needle even more challenging. During day time, Fig. 9 shows that the predicted airspeed trajectory accurately follows the FDR sensory reading. The random spikes observed throughout the flight represent cases when the gauge was occluded by the pilot or co-pilot. Figure 8 provides few examples of such occlusion scenarios occurring at different instances throughout the flight. Figure 10 presents the predicted trajectory vs. the FDR readings of N1/NG1 sensor values. The few random spikes in the prediction are due to occlusion by the pilot or copilot head or hand movements.

Table 2. Gauges Validations Accuracy

Gauge	Bin Size(α)	Validation Accuracy
Airspeed (kn)	1(kn)	78.4%
Airspeed (kn)	3(kn)	87.4%
Airspeed (kn)	5(kn)	92.1%
N1 Gauge 1	1	89%
N1 Gauge 2	1	89%
Tachometer	1	88%

CONCLUSION AND FUTURE WORK

In this paper, we proposed a video analytics framework based on deep learning to automate retrieval of

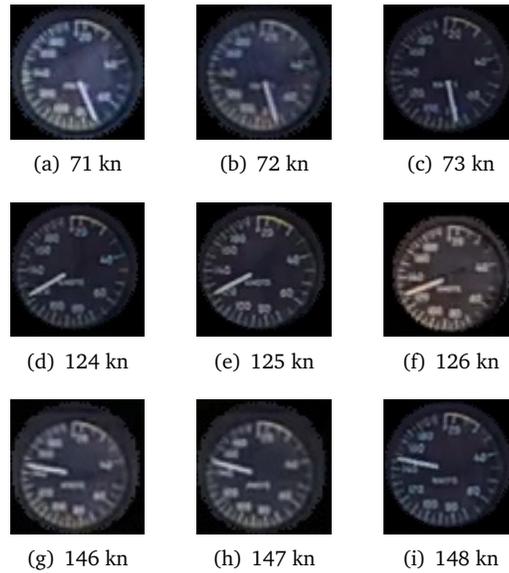


Fig. 6. The first row presents images of airspeed gauges with 71 kn, 72 kn and 73 kn. The second row presents the airspeed gauges with 124 kn, 125 kn and 126 kn. The third row presents the airspeed gauges with 146 kn, 147 kn and 148 kn. The analog needle position looks the whereas the FDR readings are different. Even the human eye cannot distinguish between consecutive airspeed.

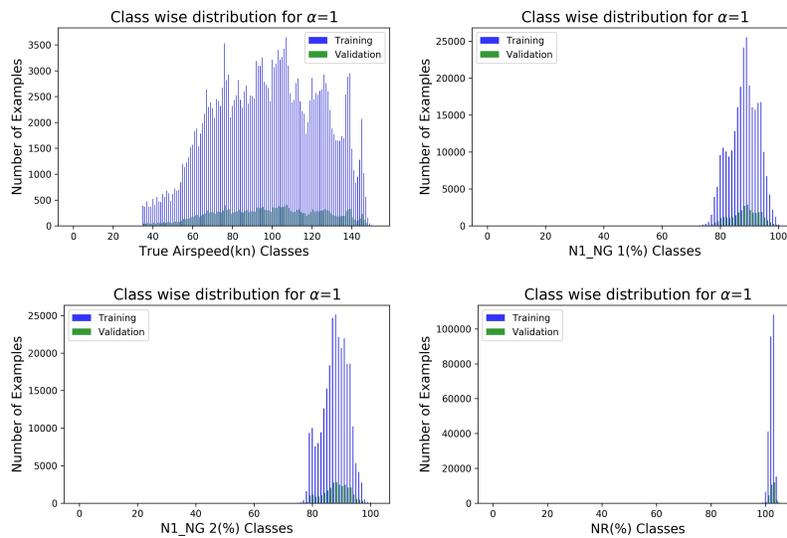
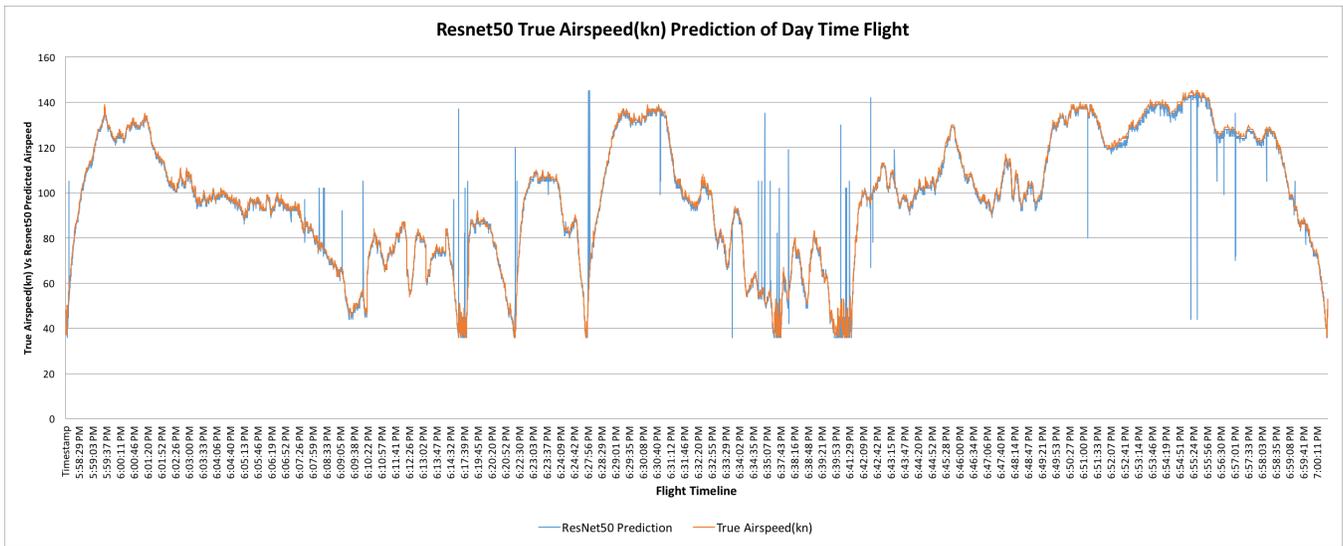


Fig. 7. A diagram presenting the class-wise data distribution of airspeed, N1/NG1, N1/NG2 and tachometer gauges.



(a) Resnet50 predicted trajectory vs FDR reading of airspeed



(b) Timestamp 5:52:59.32 PM



(c) Timestamp 6:07:39.67 PM



(d) Timestamp 6:09:09.47 PM



(e) Timestamp 6:26:27.82 PM



(f) Timestamp 6:30:47.55 PM



(g) Timestamp 6:42:26.30 PM



(h) Timestamp 6:42:27.63 PM

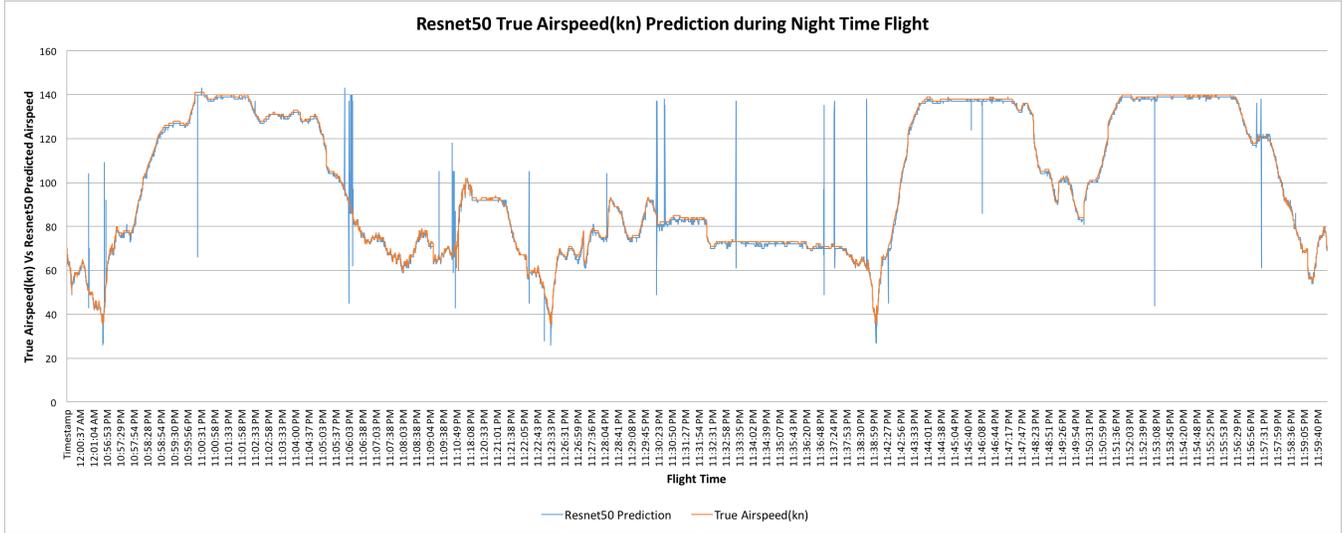


(i) Timestamp 6:56:02.65 PM



(j) Timestamp 6:58:12.58 PM

Fig. 8. The sub-figure 8(a) presents the Resnet50 predicted trajectory of a day light flight video vs. FDR recorded values for the airspeed gauge. The random predictions (spikes) are caused when the gauge was occluded due to the pilot or copilot head or hand movements. Sub-figure 8(b) shows an unobstructed camera view (without occlusion). Sub-figures 8(c) to 8(j) present frames when occlusion occurred.



(a) Resnet50 predicted trajectory vs FDR reading of airspeed



(b) Timestamp 12:00:11.67 AM



(c) Timestamp 12:00:24.20 AM



(d) Timestamp 10:06:08.27 PM



(e) Timestamp:11:30:35.02 PM



(f) Timestamp:11:30:35.02 PM



(g) Timestamp:11:38:44.16 PM



(h) Timestamp:11:42:29.43 PM



(i) Timestamp:11:53:03.97 PM



(j) Timestamp:11:57:07.70 PM

Fig. 9. The sub-figure 9(a) presents the Resnet50 predicted trajectory at night vs. FDR recorded values for airspeed gauge. Sub-figure 9(b) shows an unobstructed camera view (without occlusion). Sub-figures 9(c) to 9(j) present frames when the gauge was occluded.

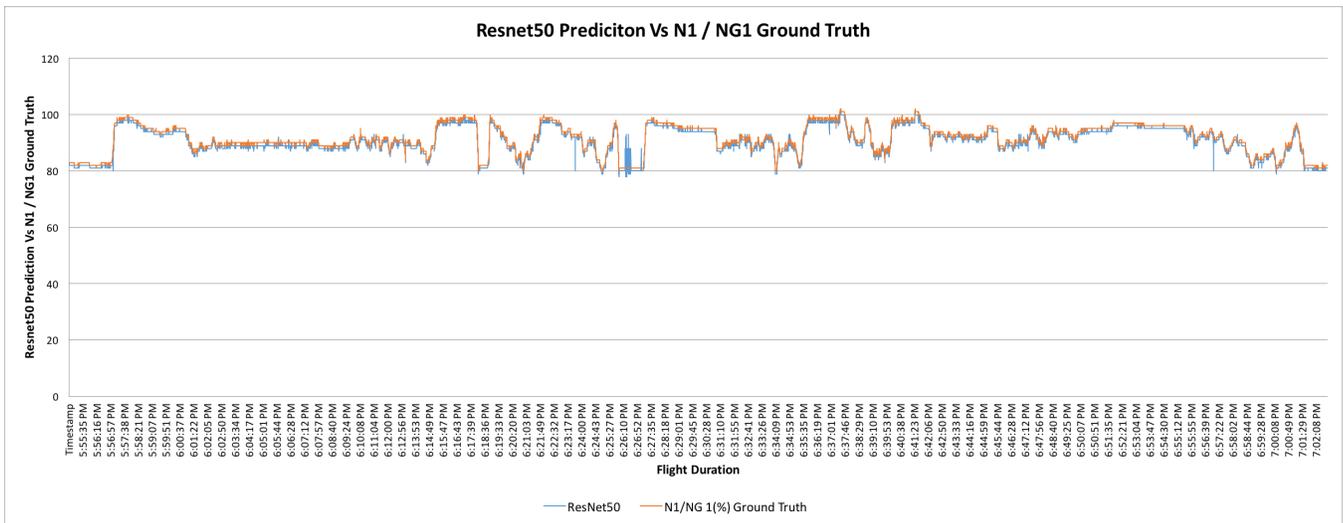


Fig. 10. Resnet50 prediction for a test flight video. The predicted trajectory closely follows the actual FDR sensor recording of N1/NG1 gauge except for few spikes due to occlusion.

flight state information from cockpit instrument panel videos. The proposed approach, based on Resnet50 model, achieved validation accuracies of 78%, 89%, 89% and 88%, for airspeed, N1/NG 1, N1/NG 2 and tachometer gauges, respectively. These results demonstrated the feasibility of an inexpensive cockpit camera solution that would facilitate participation of rotorcraft industry in FDM programs. Such optical sensors will result in an effective flight data analysis tool as well as improved safety and operational efficiency of the rotorcraft. The proposed approach also speeds up the post flight video data analysis and offers a cost effective method for retrofitting older rotorcraft with FDR-like devices. In the future, we will extend the scope of this work to estimate the other FDR sensor readings using flight videos. We will also include a pre-processing step that will increase the resolution of the gauges prior to network training. The pre-processing step will also be based on a super-resolution CNN. The hope is that enhanced resolutions input frames would increase the performance accuracy of the network. Future work also includes a hybrid model architecture of a CNN and a Long Term Short Memory (LSTM) network to address the challenging scenarios, such as pilot obstruction of the instrument panel, sun light distortion and night time conditions. LSTMs account for temporal correlations between the frames, and thus could improve accuracy even during occlusion.

ACKNOWLEDGMENTS

This work was supported by the Federal Aviation Administration (FAA) under Grant number 16-G-015 and the National Science Foundation under Grant NSF DUE-1610911.

REFERENCES

- ¹United States Helicopter Safety Team, Helicopter Facts, Accident Rates Differ Between Helicopter and Fixed-Wing Aircraft, http://www.ihst.org/portals/54/newsletters/Facts4_Accident%20Rates.pdf
- ²United States Department of Transportation, Federal Aviation Administration, Revisions to Cockpit Voice Recorder and Digital Flight Data Recorder Regulations, RIN 2120-AH88, https://www.faa.gov/regulations_policies/rulemaking/recently_published/media/23532.DOC
- ³Braman, G., “History of System Safety,” American Helicopter Society 74th Annual Forum, Phoenix, Arizona, USA, May 2018.
- ⁴National Transportation Safety Board, Safety Recommendation A-00-031 https://www.nts.gov/_layouts/nts.recsearch/Recommendation.aspx?Rec=A-00-031
- ⁵Cooke, M., Webb, T., Kramer, M., Edel, Z., and Packard, C., “Rotary Inspection Tracking System,” American Helicopter Society 74th Annual Forum, Phoenix, Arizona, USA, May 2018.
- ⁶Hewitt, J., and Foito, D., “Advantages of Early System Safety Influence,” American Helicopter Society 74th Annual Forum, Phoenix, Arizona, USA, May 2018.
- ⁷United States Department of Transportation, Federal Aviation Administration, System Safety Handbook https://www.faa.gov/regulations_policies/handbooks_manuals/aviation/risk_management/ss_handbook/

- ⁸Continuing Decrease for Helicopter Accidents in Key Regions Around the World <http://www.ihst.org/Default.aspx?tabid=1507&mid=2918&newsid2918=64036&language=en-US>
- ⁹IHST Global Survey Underscores Safety Strengths and Weaknesses in the Helicopter Industry <http://www.ihst.org/Default.aspx?tabid=1507&mid=2918&newsid2918=64037&language=en-US>
- ¹⁰U.S. Team Releases Details on Proposed Safety Enhancements Focused on Reducing Fatal Helicopter Accidents <http://www.ihst.org/Default.aspx?tabid=1507&mid=2918&newsid2918=64033&language=en-US>
- ¹¹Shin, S., and Hwang, I., "Helicopter Cockpit Video Data Analysis for Attitude Estimation using DBSCAN Clustering", AIAA Infotech at Aerospace, San Diego, California, USA 2016, doi:10.2514/6.2016-0920.
- ¹²Kuo, B. C., Guan, W., Chen, P., Hsiao, F., and Chang, F., "Helicopter Lightweight Flight Recorder Image Analysis for Flight Data Monitoring Purpose," Integration, and Operations Conference, AIAA AVIATION Forum, 2018, doi:10.2514/6.2018-3051.
- ¹³Hu, J., Shen, L., Albanie, S., Sun, G., and Wu, E., "Squeeze-and-excitation networks," *IEEE Conference on Computer Vision and Pattern Recognition* 2018.
- ¹⁴Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., and Yuille, A. L., "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol 4 May 2018, pp. 834-848. doi:10.1109/TPAMI.2017.2699184
- ¹⁵Redmon, J., and Farhadi, A., "YOLO9000: better, faster, stronger," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* Honolulu, HI, USA, Jan 2017, pp. 834-848. doi:10.1109/CVPR.2017.690
- ¹⁶Smith, R., "An overview of the Tesseract OCR engine," Ninth International Conference on Document Analysis and Recognition, Parana, Brazil, Sept 2007, doi:10.1109/ICDAR.2007.4376991
- ¹⁷Ian, G, Yoshua, B, Aaron, C: Deep learning (Vol. 1). Cambridge: MIT press, (2016)
- ¹⁸Karen, S., and Zisserman, A., "Very deep convolutional networks for large-scale image recognition," International Conference on Learning Representations (ICLR), San Diego, USA May 2015.
- ¹⁹Pan, S. J., and Yang, Q., "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering* Vol 22, Oct 2010, pp.1345-1359. doi:10.1109/TKDE.2009.191
- ²⁰Diederik, P., and Kingma and Jimmy, B., "Adam: A Method for Stochastic Optimization," *International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, May 7-9, 2015.
- ²¹Kaiming, H., Xiangyu, Z., Shaoqing, Ren., and Jian, Sun., "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.770-778, 2016.
- ²²Francois, C., "Xception: Deep Learning With Depthwise Separable Convolutions," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1251-1258, 2017.
- ²³Christian, S., Vanhoucke, V., Sergey, I., Jonathon, S., and Zbigniew, W., "Rethinking the Inception Architecture for Computer Vision," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818-2826, 2016.