Abstract—Stroke is a leading cause of death and disability. Early prediction of stroke patients’ functional outcomes is helpful for treatments. However, current diagnosis machines, such as computed tomography (CT) and magnetic resonance imaging (MRI), are expensive, not portable, and may cause side effects. Additionally, current diagnosis scales, such as National Institutes of Health Stroke Scale (NIHSS), should be evaluated by professional medical staff, and thus cannot be conducted continuously. In this paper, we propose a multi-modal analysis methodology to predict a stroke patient’s functional outcome based on physiological signals, including EKG, ABP, and PPG. By applying the multi-modal framework to analyze the stroke patients’ physiological signals in intensive care unit (ICU), we find that the accuracy of stroke outcome predictions achieves 82.7%, which performs better than a single-modal built by any single phase. In addition, the joint EKG-ABP-PPG analysis achieves performance comparable to NIHSS, implying that the multi-modal analysis framework has potential for predicting functional outcomes of stroke.

Keywords—multi-modal; physiological signal; stroke

I. INTRODUCTION

Stroke is a leading cause of death and disability. Early prediction of functional outcomes is helpful for stroke patients’ treatments. Nowadays, two of the diagnosis machines for stroke are computed tomography (CT) and magnetic resonance imaging (MRI). CT and MRI have the advantages of precisely identifying types, locations, and outcomes. However, those diagnosis machines suffer several disadvantages. For example, the prices are high, the machines are not portable, and the radiation might cause side effects.

Currently, one of the diagnosis scales for stroke is National Institutes of Health Stroke Scale (NIHSS). NIHSS grades an overall degree of neurologic impairment, and it is one of the most reliable and valid instruments of clinical measurement in stroke [1]. However, such a clinical diagnosis should be evaluated by professional medical staff, and thus cannot be conducted continuously.

In contrast to the above-mentioned diagnosis machines and scales, physiological signals of stroke patients are inexpensive, convenient, and safe. They can also be measured continuously, even for several days. Electrocardiography (EKG), arterial blood pressure (ABP), and photoplethysmography (PPG) are common physiological signals measured by vital sign monitors. If we consider a heart to be the source of signals, and fingers to be the destinations of signals, the measurement of EKG is closer to the source, whereas the measurements of ABP and PPG are closer to the destination. Then, joint consideration among EKG, ABP, and PPG may contain the information from sources, destinations, and the channels among them. However, the ways to properly combine different phases of information to accurately predict the functional outcomes of stroke patients are still unknown.

II. PROPOSED MULTI-MODAL ANALYSIS FRAMEWORK

The flow chart of the proposed multi-modal analysis framework for stroke outcome prediction is shown in Fig. 2. Inputs of the framework are physiological signals, including EKG, ABP, and PPG. Outputs of the framework are the predictions of severity variations.

The rest of the paper is organized as follows. Section II presents the proposed multi-modal analysis framework for stroke outcome prediction. Section III presents the experimental setups and results. Section IV is the conclusion.

A. Pre-processing

The first stage of the pre-processing is to check the integrity of physiological signals. All input signals are divided into 60-second segments. A segment with severe motion artifacts, high frequency
where \( m \) is the pattern length and \( r \) is the similarity criterion. The probability \( B^m(r) \) measures whether the distance between two vectors is lower than \( r \) in \( m \)-dimensional space. Thus, the more complex the signal is, the higher the entropy will be. However, the measured signal consists of not only the physiological signal coming from the subject, but also the additional “trend” caused by the background signals. The additional trend leads to underestimate \( \text{SampEn} \). As shown in Fig. 3, after adding an additional trend to \( 1/f \) noise, \( \text{SampEn} \) on each scale decreases. Therefore, empirical mode decomposition (EMD) [7] is applied to de-trend before the calculation of MSE.

\[ \text{AUC} = \sum_{i=1}^{20} \text{SampEn}(y^i). \]  

(3)

\( \text{AUC} \) reflects the overall complexity of a person from short-term regulation to long-term regulation.

For feature transformation, there exist some algorithms, such as dimension reduction and linear discriminant analysis. However, in order to easily explain the physiological meanings of transformed MSE features, simple transformation methods are preferred. Therefore, the values of 20 MSE scales are also transformed into area under the curve (AUC) by summing them up. That is:

C. Feature Transformation and Feature Selection

Feature transformation and selection are used to transform features and select dominant features in our proposed framework, because the number of samples (i.e., stroke patients) is limited when compared to the number of extracted features. Lots of features are extracted from various physiological signals and various domains. There are 26, 104, and 104 features extracted from EKG, ABP, and PPG, respectively. These features might be effective in stroke outcome prediction. However, inputting all of the extracted features into a classifier cannot achieve acceptable performance due to some problems which result from limited number of samples, such as “curse of dimensionality” and over-fitting. Therefore, feature transformation and selection are necessary to (a) avoid over-fitting and improve model performance, (b) provide cost-effective models, and (c) gain a deeper insight into the underlying process that generates the data. [8]

For feature transformation, there exist some algorithms, such as principal component analysis and linear discriminant analysis. However, in order to easily explain the physiological meanings of transformed MSE features, simple transformation methods are preferred. Therefore, the values of 20 MSE scales are also transformed into area under the curve (AUC) by summing them up. That is:

\[ \text{AUC} = \sum_{i=1}^{20} \text{SampEn}(y^i). \]  

(3)

\( \text{AUC} \) reflects the overall complexity of a person from short-term regulation to long-term regulation.

For feature selection, there are also some algorithms, such as analysis of variance (ANOVA), information gain, Euclidean distance, and recursive feature selection (RFS). In this paper, RFS is adopted for feature selection. Besides selecting the features which can achieve higher predictive accuracy, RFS considers the dependency among features, which implies that a variable which is useless by itself can be useful when it is considered together with others. In general, RFS recursively eliminates the entire feature set into a smaller subset by considering the coefficients of a linear model. The steps of RFS are illustrated in Table II. By the way, only one of the ABP parameters and one of the PPG parameters are selected after comparing the performances of various joint EKG-ABP-PPG analyses.

D. Classification

ANOVA is a widely used statistical method which enables us to judge whether the differences among groups are significant or not. However, through such a statistical method, we still cannot tell whether a new patient’s functional outcome is good or bad. Therefore, techniques for data classification should be adopted.

There are different techniques for data classification, such as the naive Bayes classifier, neural network, and support vector machine (SVM). Linear kernel SVM is adopted to build the multi-modal framework for stroke outcome prediction considering the number and the properties of our samples. Moreover, stratified three-fold cross-validation is applied to avoid over-fitting. The steps of classification are illustrated in Table III. The flow chart of feature transformation, feature selection, and classification is shown in Fig. 4.
TABLE II. RECURSIVE FEATURE SELECTION (RFS) [9-10]

Input
Assume there are K samples. After the process of feature extraction, there are N features for each sample, \( x_j = [x_{j1}, x_{j2}, ..., x_{jn}] \), \( 1 \leq j \leq K \).
The features of all samples form an array, \( X = [x_1, x_2, ..., x_K] \).
Class labels of the samples:
\( y = [y_1, y_2, ..., y_K] \).

Initialize
Subset of the surviving features:
\( s = [s_1, s_2, ..., s_K] \).
Ranked list of the eliminating features:
\( r = [] \).

Repeat the following steps until \( s = [] \):
S1. Restrict the features of the training samples to the subset of the surviving features:
\( X' = X(s,:) \).
S2. Train the classifier:
\( SVM_{train}(X', y) \).
S3. Compute the weight vector:
\( w \).
S4. Compute the ranking criteria:
\( c = w^T w \).
S5. Find the feature with the smallest ranking criterion:
\( c_{min} = \text{argmin}(c) \).
S6. Update the ranked list of eliminating features:
\( r = [s(c_{min}), r] \).
S7. Eliminate the feature with the smallest ranking criterion:
\( s = s(1: c_{min} - 1, c_{min} + 1:length(s)) \).

Output
Feature ranked list, \( r \).

TABLE III. CLASSIFICATION

Input
Assume there are K samples. After the process of feature selection, there are M features for each sample, \( x_j = [x_{j1}, x_{j2}, ..., x_{jm}] \), \( 1 \leq j \leq K \).
The features of all samples form an array, \( X = [x_1, x_2, ..., x_K] \).
Class labels of the samples:
\( y = [y_1, y_2, ..., y_K] \).

Steps
S1. Conduct scaling on the selected features
S2. Consider the linear kernel support vector machine (SVM)
S3. Apply stratified three-fold cross-validation
S4. Evaluate the proposed multi-modal framework by accuracy, precision, recall, and receiver operating characteristic (ROC)

III. EXPERIMENTAL SETUPS AND RESULTS

A. Experimental Setups
The experimental data come from the intensive care unit (ICU) in National Taiwan University Hospital. One-hour physiological signals, including EKG, ABP, and PPG, are collected simultaneously by vital sign monitors within 24 hours after a stroke patient is admitted to the hospital. The sampling frequencies of EKG, ABP, and PPG are 512 Hz, 128 Hz, and 128 Hz, respectively. Besides, only the stroke patients without atrial fibrillation (AF) are involved in the analysis because the heart rates of AF stroke patients are arrhythmia.

To evaluate the accuracy of our stroke outcome prediction, a clinical scale, modified Rankin Scale (mRS), is adopted. The mRS grades the functional outcomes of stroke patients, ranging from mRS 0-no symptoms to mRS 6-death [1]. In our experiments, mRS 0-2 and mRS 3-6 are grouped into good functional outcome and bad functional outcome, respectively. The total number of analyzed patients is 75. Thirty-two of them are mRS 0-2, and forty-three of them are mRS 3-6.

B. Results of Feature Extraction and Statistical Analysis
ANOVA is adopted for statistical analysis. ANOVA calculates \( p \)-values by comparing the relative values between the variation within groups and the variation among groups. A common threshold for significant statistical difference is 0.05. The \( p \)-values of features extracted from the EKG parameter, RR, are shown in Fig. 5. We can see that \( AUC \), which is extracted by MSE analysis, outperforms traditional features.

After comparing the performances of various joint EKG-ABP-PPG analyzes, the features extracted from RR (the EKG parameter), PP (one of the ABP parameters), and AMP (one of the PPG parameters) are regarded as the inputs of RFS to build the final multimodal framework. Fig. 6 shows the MSE curves of RR, PP, and AMP, which are consistent with the physiological control mechanism. That is, the \( SampEn \) of the non-AF stroke patients with mRS 0-2 is higher than that of the non-AF stroke patients with mRS 3-6 on every scale (for RR) or on scales with large \( r \) (for PP and AMP).

Fig. 5. The \( p \)-values of features extracted from the EKG parameter, RR. The common threshold for significant statistical difference is represented by the red dot line. \( AUC \) is far below the red dot line compared with other features.
Fig. 6. The MSE curves of (a) RR (the EKG parameter), (b) PP (one of the ABP parameters), and (c) AMP (one of the PPG parameters). Stroke patients with mRS 0-2 and mRS 3-6 are respectively represented by the blue line and red line. Values are expressed as mean ± standard error of the mean. *" indicates significant statistical difference.
The performance comparisons among the multi-modal framework and three single-modal frameworks are shown in Table IV and Fig. 8. They are evaluated by accuracy, precision, recall, and receiver operating characteristic (ROC) curves. The multi-modal framework achieves the accuracy of 82.7%, precision of 84.2%, and recall of 86.0%, which are higher than any single-modal framework.

The existence of wrong prediction is reasonable and acceptable because there exist some other factors, such as age, sex, subtype of stroke (ischemic or hemorrhagic), and location of stroke (cortical, subcortical, or brainstem). Besides, even without taking these factors into consideration, the proposed multi-modal framework has already achieved performance comparable to the accuracy (77.3%), precision (83.6%), and recall (76.7%) of NIHSS, implying that the framework has potential for predicting the functional outcomes of stroke.

IV. CONCLUSION

In this paper, the multi-modal analysis methodology is proposed to predict a stroke patient’s functional outcome based on different phases of physiological signals. The physiological signals are inexpensive, convenient, and safe. They can also be measured continuously. Furthermore, by using the multi-modal framework of EKG, ABP, and PPG to analyze the stroke patients in ICU, we find that the accuracy, precision, and recall of stroke outcome predictions achieves 82.7%, 84.2%, and 86.0%, respectively. The joint EKG-ABP-PPG analysis not only performs better than a single-modal framework built by any single phase, but also achieves performance comparable to NIHSS, implying that the framework has potential for predicting the functional outcomes of stroke.

ACKNOWLEDGMENT

This work has been supported by: (a) National Taiwan University (NTU)-National Taiwan University Hospital (NTUH)-MediaTek Innovative Medical Electronics Research Center, NTUH PC851. (b) NOVATEK Fellowship.

REFERENCES


---

**TABLE IV. PERFORMANCE COMPARISON AMONG THE MULTI-MODAL FRAMEWORK AND THREE SINGLE-MODAL FRAMEWORKS**

<table>
<thead>
<tr>
<th>Framework</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-modal framework of EKG</td>
<td>70.7%</td>
<td>73.1%</td>
<td>76.7%</td>
</tr>
<tr>
<td>Single-modal framework of ABP</td>
<td>68.0%</td>
<td>72.6%</td>
<td>69.4%</td>
</tr>
<tr>
<td>Single-modal framework of PPG</td>
<td>68.0%</td>
<td>73.1%</td>
<td>69.7%</td>
</tr>
<tr>
<td>Multi-modal framework (joint EKG-ABP-PPG analysis)</td>
<td>82.7%</td>
<td>84.2%</td>
<td>86.0%</td>
</tr>
</tbody>
</table>

*True positive, TP; true negative, TN; false positive, FP; false negative, FN.*

**Fig. 7. Flow chart of the single-modal frameworks of EKG, ABP, and PPG.**

**Fig. 8. A comparison of ROC curves:** (a) the multi-modal framework with feature transformation and selection vs. without feature transformation and selection (b) the multi-modal framework vs. three single-modal frameworks.

**C. Results of Feature Transformation, Selection, and Classification**

In our experiment, the single-modal frameworks of EKG, ABP, and PPG are built to illustrate that the proposed multi-modal framework performs better than all of the single-modal frameworks. These single-modal frameworks are based on the procedure similar to Fig. 2, as shown in Fig. 7.

Compared to the single-modal frameworks, the features selected for the multi-modal framework of EKG, ABP, and PPG contain all of the input phases. That means all of the phases contribute to the building of the multi-modal framework, and therefore the decision of multi-modal framework can be regarded as the result of their joint decision. In addition, the advantages of RFS mentioned in Section II are verified: (a) A variable useless by itself can be useful when considered together with others. That is, the multi-modal framework takes several variability features into consideration, even though few variability features show significant statistical difference. (b) A cost-effective model is provided by reducing the number of features to 71.4%.

To illustrate that the procedures of feature transformation and feature selection are necessary for the multi-modal framework, another experiment is set up by removing these procedures in Fig. 2. The result is shown in Fig. 8 (a), which reveals that the assistance of feature transformation and feature selection improves the area under the ROC curve by 1.4 times. That is because reducing the dimension avoids over-fitting.